

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2002, 16:28:59 ; Search time 1539.98 seconds

(Without alignments)
22389.320 Million cell updates/sec

Title: US-09-636-826-1

Perfect score: 2090

Sequence: 1 GATATCACACATTCGTCAT.....TCGCTTCTCGGTCTGCTGCT

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1472140 seqs, 8248589755 residues

Word size: 0

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: listing first 45 summaries

Database:

GenEmbl:*
1: gb_da:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_hgtg_hum:*
31: em_hgtg_inv:*
32: em_hgtg_rnd:*
33: em_hgtg_hum:*
34: em_hgtg_inv:*
35: em_hgtg_rnd:*
36: em_hgtg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2090	100.0	2090	6 AX088250	AX088250 Sequence
2	29	1.4	991	8 MCVTRPFC	X9499 M. crystalli
3	26	1.2	13862	3 CEF39B1	Z69660 Caenorhabdi
4	25	1.2	86156	2 HUAC002550	AC002550 Human Chr
5	25	1.2	166116	2 AL360221	AL360221 Homo sapi
6	25	1.2	184666	2 AC013508	AC013508 Homo sapi
7	25	1.2	198120	2 AC032169	AL392169 Homo sapi
8	25	1.2	193305	2 AC093278	AC093278 Homo sapi
9	25	1.2	199517	2 AL353741	AL353741 Human DNA
10	25	1.2	211271	2 AL590624	AL590624 Homo sapi
11	24	1.1	1366	3 PFCBPH21	X6976 P. falciparu
12	24	1.1	8077	2 AC019982	X79836 P. falciparu
13	24	1.1	107257	10 AF289665	AC019982 Drosophila
14	24	1.1	114898	10 AF139987	AF289665 Mus muscu
15	24	1.1	123004	2 AC014962	AF139987 Mus muscu
16	24	1.1	148744	9 AC012636	AC014962 Drosophila
17	24	1.1	175561	9 AC012636	AL355576 Human DNA
18	24	1.1	177129	2 AC037481	AC012636 Homo sapi
19	24	1.1	179947	2 AC009786	AC037481 Homo sapi
20	24	1.1	192581	2 PFMA1351	AC009786 Homo sapi
21	24	1.1	194188	2 AC044857	AL049180 Plasmodiu
22	24	1.1	200849	2 AC091250	AC044857 Homo sapi
23	24	1.1	224843	2 AC021223	AC091250 Mus muscu
24	24	1.1	289887	3 AE003542	AC021223 Homo sapi
25	24	1.1	300503	2 AC010260	AE003542 Drosophila
26	24	1.1	305018	3 AE003437	AC010260 Homo sapi
27	24	1.1	1208	5 AF123885	AE003437 Drosophila
28	23	1.1	38192	2 AC020177	AF123885 Gallus ga
29	23	1.1	61756	2 AC017569	AC020177 Drosophila
30	23	1.1	75762	2 AC036137	AC017569 Drosophila
31	23	1.1	77585	2 AL133418	AC036137 Homo sapi
32	23	1.1	77635	2 DMBR4001	AL133418 Human DNA
33	23	1.1	90157	2 AL355552	AL122024 Drosophila
34	23	1.1	103675	2 HSDJ93P18	AL355552 Human DNA
35	23	1.1	113733	2 AP003159	AL080251 Homo sapi
36	23	1.1	11674	2 DMBR30C13	AP003159 Homo sapi
37	23	1.1	130494	2 AL139422	AL122025 Drosophila
38	23	1.1	133221	2 HS339A15	AL139422 Homo sapi
39	23	1.1	141184	2 AL139283	AL031423 Human DNA
40	23	1.1	144260	9 AC006998	AL139283 Homo sapi
41	23	1.1	146516	2 AC079981	AC006998 Homo sapi
42	23	1.1	148430	9 AC041034	AC079981 Homo sapi
43	23	1.1	153724	2 AC041034	AL139328 Human DNA
44	23	1.1	154935	2 AL356781	AC041034 Homo sapi
45	23	1.1	154935	2 AL356781	AL356781 Homo sapi

ALIGNMENTS

RESULT 1	AX088250	2090 bp	DNA	PAT	17-MAR-2001
LOCUS	AX088250	Sequence 1 from Patent WO0114572.			
DEFINITION	AX088250				
ACCESSION	AX088250				
VERSION	AX088250.1	GI:13397159			
KEYWORDS					
SOURCE	Beta vulgaris.				
ORGANISM	Beta vulgaris				
REFERENCE					
AUTHORS	Duvenig, E. and Rausch, T.				
TITLE	Plant gene expression, controlled by constitutive plant v-atpase promoters				
JOURNAL	Patent: WO 0114572-A 1 01-MAR-2001;				
FEATURES	BASE AKTINGESELISCHAFT (DE)				
source	Location/Qualifiers				
	1..2090				

/organism="Beta vulgaris"
/db_xref="taxon:161934"
/note="Promotor subunit c isoform 2"
1..1923
BASE COUNT 794 a 354 c 271 g 671 t
ORIGIN

```
Query Match      100.0%; Score 2090; DB 6; Length 2090;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2090; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GATATCACACATTCCTCCATCGACGATTTGGCACTTCAAAATAGTACGTAATCTTTT 60  
DB 1 GATATCACACATTCCTCCATCGACGATTTGGCACTTCAAAATAGTACGTAATCTTTT 60  
QY 61 AATCTTCAAAAGTATTTTCACATTTAGTACGATTTGCAACTTCAAAATAGTACGTAAT 120  
DB 61 AATCTTCAAAAGTATTTTCACATTTAGTACGATTTGCAACTTCAAAATAGTACGTAAT 120  
QY 121 GGAATTAAGTGAAGAGGTGTTATATATATAGACTTAAATTTGATTCATTTTCATATAT 180  
DB 121 GGAATTAAGTGAAGAGGTGTTATATATATAGACTTAAATTTGATTCATTTTCATATAT 180  
QY 181 CTGAAGACAGGATGATGATGAATTTGATTCATTTATGACATGATGAAGATTAACGA 240  
DB 181 CTGAAGACAGGATGATGATGAATTTGATTCATTTATGACATGATGAAGATTAACGA 240  
QY 241 TTGATCTCTTTTAAATTCATATATATATTTTGGCCAAATCTTTGCAAAATATC 300  
DB 241 TTGATCTCTTTTAAATTCATATATATATTTTGGCCAAATCTTTGCAAAATATC 300  
QY 301 CATGTGCGAAATATATTTGAAACCAACATATCAACCTTTTGGCAACACTTT 360  
DB 301 CATGTGCGAAATATATTTGAAACCAACATATCAACCTTTTGGCAACACTTT 360  
QY 361 TACAAAATTCATTTTCAGAAAAAAATTTTACATTAACCTTGCAGAAATCAAAATGTTG 420  
DB 361 TACAAAATTCATTTTCAGAAAAAAATTTTACATTAACCTTGCAGAAATCAAAATGTTG 420  
QY 421 GAAAAATTTAAATTTCTCTTCACTATATATGAAAGTCAAGTGAATTAATTAAGAAA 480  
DB 421 GAAAAATTTAAATTTCTCTTCACTATATATGAAAGTCAAGTGAATTAATTAAGAAA 480  
QY 481 GGAGAAAAATAAAAATGACATTTTCATGCGAAATCAAAATGCTATGAAAAACTTTAAA 540  
DB 481 GGAGAAAAATAAAAATGACATTTTCATGCGAAATCAAAATGCTATGAAAAACTTTAAA 540  
QY 541 TTTTATTTTAAATATATATGAAATCAAGTGTAAATTTAGAAAAAGGAAAAATTA 600  
DB 541 TTTTATTTTAAATATATATGAAATCAAGTGTAAATTTAGAAAAAGGAAAAATTA 600  
QY 601 AATGACCATTTTCATCAAAATCAGATGTGTATGAAAAATTTAAATTTTATTTCAATA 660  
DB 601 AATGACCATTTTCATCAAAATCAGATGTGTATGAAAAATTTAAATTTTATTTCAATA 660  
QY 661 TAATTTGAAGCTCAAGGTGAACATTTTGAAGAGGAAAAATTTAAATGATGAATTTT 720  
DB 661 TAATTTGAAGCTCAAGGTGAACATTTTGAAGAGGAAAAATTTAAATGATGAATTTT 720  
QY 721 GTAAACATCATTTTGTGAAATCAGATTTTGAAGTTGACAGAGAAAAAAATGTAAT 780  
DB 721 GTAAACATCATTTTGTGAAATCAGATTTTGAAGTTGACAGAGAAAAAAATGTAAT 780  
QY 781 GTCTTATCTTTTCGGTTCACAAATTTTGGGATCATTAAGAAATTTCTGAAATCCATAT 840  
DB 781 GTCTTATCTTTTCGGTTCACAAATTTTGGGATCATTAAGAAATTTCTGAAATCCATAT 840  
QY 841 AAACATTTTAAATTAACAAAAATGAATTAACCAAAAAAAGAAAGAAAGTGCATTTTC 900  
DB 841 AAACATTTTAAATTAACAAAAATGAATTAACCAAAAAAAGAAAGAAAGTGCATTTTC 900  
QY 901 GTAAGACATCATCTGATTTATTAAGAACATGCGCATATTAGAATTTGAGAAAAA 960
```

```
DB 901 GTAAGACATCATCTGATTTATTAAGAACATGCGCATATTAGAATTTGAGAAAAA 960  
QY 961 ACTATTTCAAAATCAGAAAAATGATTAACAACATCAAGAAAGATCAAAATCTTATTTCA 1020  
DB 961 ACTATTTCAAAATCAGAAAAATGATTAACAACATCAAGAAAGATCAAAATCTTATTTCA 1020  
QY 1021 CAAAATGAGGTGACTTTAAATACATTTGCAATTTTCAGTTTATTTACTACTTACTT 1080  
DB 1021 CAAAATGAGGTGACTTTAAATACATTTGCAATTTTCAGTTTATTTACTACTTACTT 1080  
QY 1081 AGGCTTAAAAATATCATGACAGATCGGCTGCACAAAAGACATGATTAAGATCACAA 1140  
DB 1081 AGGCTTAAAAATATCATGACAGATCGGCTGCACAAAAGACATGATTAAGATCACAA 1140  
QY 1141 GTCAACAACTCACAACATGATTTTCATTTAGGCTCATTTGTTGAGGCGTAAACGTTTC 1200  
DB 1141 GTCAACAACTCACAACATGATTTTCATTTAGGCTCATTTGTTGAGGCGTAAACGTTTC 1200  
QY 1201 CCGGACACATTTTCTCCATTTTATGATTTACATTTGTTGAGGCGTAAACGTTTC 1260  
DB 1201 CCGGACACATTTTCTCCATTTTATGATTTACATTTGTTGAGGCGTAAACGTTTC 1260  
QY 1261 AACGTTTTCCTTGGGGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1320  
DB 1261 AACGTTTTCCTTGGGGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1320  
QY 1321 AATGAAGGAAACATGTTTCTCTATCTCTCTGTTACACTTTTCCATCTCTCTTAC 1380  
DB 1321 AATGAAGGAAACATGTTTCTCTATCTCTCTGTTACACTTTTCCATCTCTCTTAC 1380  
QY 1381 TTTCCCTTTTATTTTACTTTTCAATTTCTTTTCTTTGATGAACCAACCAACGGA 1440  
DB 1381 TTTCCCTTTTATTTTACTTTTCAATTTCTTTTCTTTGATGAACCAACCAACGGA 1440  
QY 1441 ACTATTTTGAATTTGTTTCCATGTAATTTTCCATGAAGAAATCAATTTTACT 1500  
DB 1441 ACTATTTTGAATTTGTTTCCATGTAATTTTCCATGAAGAAATCAATTTTACT 1500  
QY 1501 GAAAAATTTTAAATTTCTCTTCACTATATATGAAAGTCAAGTGAATTAATTAAGAAA 1560  
DB 1501 GAAAAATTTTAAATTTCTCTTCACTATATATGAAAGTCAAGTGAATTAATTAAGAAA 1560  
QY 1561 CAATTTGCAAACTGTTTACCGTTTCCATTTGTTTGGATTTGATTAAGAAAGTATTC 1620  
DB 1561 CAATTTGCAAACTGTTTACCGTTTCCATTTGTTTGGATTTGATTAAGAAAGTATTC 1620  
QY 1621 AACTATTTGCTCACACCTTAAGATGATTAAGTAAACCTTGAAGTGGGCTTTGGCTAAA 1680  
DB 1621 AACTATTTGCTCACACCTTAAGATGATTAAGTAAACCTTGAAGTGGGCTTTGGCTAAA 1680  
QY 1681 AAAAGTCCACCAAGCCCAATTTAGGCTTCCCAAAACCGAAATTTCTGCTACTAT 1740  
DB 1681 AAAAGTCCACCAAGCCCAATTTAGGCTTCCCAAAACCGAAATTTCTGCTACTAT 1740  
QY 1741 CCMAAAGCAAAAGCAACCTCTGATCAACCGAAATTAACATATTTTGTGTTT 1800  
DB 1741 CCMAAAGCAAAAGCAACCTCTGATCAACCGAAATTAACATATTTTGTGTTT 1800  
QY 1801 CTCCCAATTTTCAATTTTATTTATCAAGGGAAGTACAAATTAATCAATCAAGGTTA 1860  
DB 1801 CTCCCAATTTTCAATTTTATTTATCAAGGGAAGTACAAATTAATCAATCAAGGTTA 1860  
QY 1861 AAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1920  
DB 1861 AAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1920  
QY 1921 CTCATATTTCAATCAAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1980  
DB 1921 CTCATATTTCAATCAAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1980  
QY 1981 TCGGATCAAGCAATTCATGCAACCTTCCGATATATACCATCAAAAAAATGTCACAG 2040
```

```

Db 1981 TCGATCAGCAATTCATCGACACCTTCGATCATCACCATCAAAAAATGTCAACAG 2040
Oy 2041 TCTTACGGCGATGAACGGCCGCTTCGCTTCGGGCTGCTGCT 2090
Db 2041 TCTTACGGCGATGAACGGCCGCTTCGCTTCGGGCTGCTGCT 2090

RESULT 2
MCVATPC 991 bp mRNA PLN 21-OCT-1996
LOCUS M.crystallinum mRNA for v-type ATPase c subunit.
ACCESSION X94999
VERSION X94999.1 GI:1495682
KEYWORDS ATPase c subunit.
SOURCE common ice plant.
ORGANISM Mesembryanthemum crystallinum
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Caryophyllales; Caryophyllales; Alzooceae; Mesembryanthemum.
REFERENCE
AUTHORS Tsiantis,M.S., Bartholomew,D.M. and Smith,J.A.
TITLE Salt regulation of transcript levels for the c subunit of a leaf
vacuolar H(+) -ATPase in the halophyte Mesembryanthemum crystallinum
JOURNAL Plant J. 9 (5), 729-736 (1996)
MEDLINE 96337458
REFERENCE
AUTHORS Tsiantis,M.S.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-1996) M.S. Tsiantis, Dept. of Plant Sciences,
University of Oxford, South Parks Rd., Oxford, OX1 3RB, UK

FEATURES
SOURCE
1. 991
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/clone_lib="Unizap-XR"
115.. 612
/gene="Vmacl1"
115.. 612
/gene="Vmacl1"
/codon_start=1
/product="V-type ATPase c subunit"
/protein_id="CAA64455.1"
/db_xref="GI:1495683"
/db_xref="SWISS-PROT:Q39437"
/translation="MSTVNGDETAFEFGLGAALVPSGGAAYGTAQSGVGYAM
GVKPELVKRSIVPYMAGVIGLIVITSTGINPKRSYLYFPDVAHSSGLAC
GIAGISGMAIGIVGDAVYRANAOQPKLVGMILLIFPAHALALGLIVGITLSRAS
OSRAD"
864.. 869
polyA_signal
973.. 978
polyA_signal
BASE COUNT 195 a 216 c 220 g 360 t
ORIGIN

Query Match 1.43; Score 29; D8 8; Length 991;
Best Local Similarity 100.0%; Pred.No. 0.026;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2049 GGCGATGAACGGCCGCTTCGCTTCGGGCTGCT 2077
Db 133 GGCGATGAACGGCCGCTTCGCTTCGGGCT 161

RESULT 3
CEP39B1 13862 bp DNA INV 25-OCT-2000
LOCUS Caenorhabditis elegans cosmid F39B1, complete sequence.
ACCESSION Z69660
VERSION Z69660.1 GI:1200273
KEYWORDS H1e; Phosphatidylinositol 3-kinase like; Transfer RNA; tRNA-Val.
SOURCE Caenorhabditis elegans.
ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;

```

```

REFERENCE
AUTHORS none.
TITLE Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium
JOURNAL Science 282 (5396), 2012-2018 (1998)
MEDLINE 99069613
REMARK The C.elegans Sequencing Consortium.
Erratum: [[published errata appear in Science 1999 Jan
1:283(5398):35 and 1999 Mar 26:283(5410):2103 and 1999 Sep
3:283(5433):1493]]
2 (bases 1 to 13862)
REFERENCE
AUTHORS Lightning,J.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-1996) Nematode Sequencing Project, Sanger Centre,
Hinxton, Cambridge CB10 1RQ, England and Department of Genetics,
Washington University, St. Louis, MO 63110, USA. E-mail:
jassanger.ac.uk or twenematode.wustl.edu
Coding sequences below are predicted from computer analysis, using
predictions from GeneFINDER (P. Green, U. Washington), and other
available information.
Current sequence finishing criteria for the C. elegans genome
sequencing consortium are that all bases are either sequenced
unambiguously on both strands, or on a single strand with both a
dye primer and dye terminator reaction, from distinct subclones.
Exceptions are indicated by an explicit note.
IMPORTANT: This sequence is not the entire insert of clone F39B1.
It may be shorter because we only sequence overlapping sections
once, or longer because we arrange for a small overlap between
neighbouring submissions.
The true left end of clone F39B1 is at 1 in this sequence. The true
right end of clone F39B1 is at 21038 in
sequence Z69903.
The true left end of clone F46f2 is at 13764 in this sequence. The
start of this sequence (1..100) overlaps with the end of sequence
Z93239.
The end of this sequence (13764..13862) overlaps with the start of
sequence Z69903.
For a graphical representation of this sequence and its analysis
see: http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?
name=F39B1.
FEATURES
SOURCE
1. 13862
/organism="Caenorhabditis elegans"
/db_xref="taxon:6239"
/chromosome="X"
/clone="F39B1"
Join(complement(269903..1:4329..4372),
complement(269903..1:4079..4228),
complement(269903..1:3599..3804),
complement(269903..1:3301..3468),
complement(269903..1:2880..3094),
complement(269903..1:1891..1999),
complement(269903..1:1618..1846),
complement(269903..1:269..399),complement(13603..13766),
complement(13432..13550),complement(12558..13006),
complement(12347..12507),complement(11854..12300),
complement(11183..11768),complement(10501..10937),
complement(10294..10450),complement(9998..10242),
complement(9715..9790),complement(9564..9664),
complement(9378..9513),complement(7366..7516),
complement(7166..7311),complement(6892..7089))
/gene="F39B1.1"
Join(complement(269903..1:4329..4372),
complement(269903..1:4079..4228),
complement(269903..1:3599..3804),
complement(269903..1:3301..3468),
complement(269903..1:2880..3094),
complement(269903..1:1891..1999),
complement(269903..1:1618..1846),
complement(269903..1:269..399),complement(13603..13766),
complement(13432..13550),complement(12558..13006),
complement(12347..12507),complement(11854..12300),
complement(11183..11768),complement(10501..10937),
complement(10294..10450),complement(9998..10242),
complement(9715..9790),complement(9564..9664),
complement(9378..9513),complement(7366..7516),
complement(7166..7311),complement(6892..7089))

```

[illegible]

FEATURES
Source

Location/Qualifiers
1. 86156
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16p12.3"
/clone="A-101F10"
complement(join(8517. .8779,10305. .10434,13638. .13702, 25645. 25837))
/gene="101F10.3"
complement(8517. .25837)
/gene="101F10.3"
complement(join(8517. .8779,10305. .10434,13638. .13702, 25645. 25837))
/gene="101F10.3"
/codon_start=1
/product="Unknown gene product"
/protein_id="AAC05807.1"
/db_xref="GI:2911264"
/translation="MAAPQIPSHIVRLKPKSCSTDSFTKRPVPTVLSARELPVSSN QVTEPSSKNLWEDICEYEAQDPPEEGYKQEPVITVAPEEMLEFGFSAEHFPV SHFTMISRPQPODKSETINPKTCKREYLETFFPVLPGMASLHQAARKEKEEVS CLAGFLYFELNLSLSDSSISWYHOVYLCMTSPGKACVGHLPSSSHIT"
join(29099. 30014. 32692. .32760,33546. .33623,36911. .37222)
/gene="101F10.1"
29099. .37222
/gene="101F10.1"
join(29904. .30014,32692. .32760,33546. .33623,36911. .37222)
/gene="101F10.1"
/codon_start=1
/product="Unknown gene product"
/protein_id="AAC05805.1"
/db_xref="GI:2911264"
/translation="MRSKKRVOPVLEBPALRRKKRRKESGVADPWKEETDLE VLEERKGNDEAIDQVRKALQEDIRSGCTEASRTKMTGTQFGOMDTAGENEED OKLFTLRMGGEKNIQSPFSRPASTIARPNMALGKKAADSLQONQGRDYDAMSWKYS RGAGLGFSTAPNKITFYIDRNASKSVKLED"
complement(join(43006. 43754,44484. .44630,52661. .52779, 61742. .61907,74833. .74922))
/gene="101F10.2"
complement(43006. .74922)
/gene="101F10.2"
complement(join(43656. 43754,44484. .44630,52661. .52779, 61742. .61907,74833. .74922))
/gene="101F10.2"
/codon_start=1
/product="Unknown gene product"
/protein_id="AAC05806.1"
/db_xref="GI:2911265"
/translation="MINIDGKMRPSEFLELQCNFFSLIIVDPHPEGLVETREL LNVIODYTWEDNSDEKIRIYTCVHLHLSAMSQETLYHIDKDSNDSYGGDSKFLAE NKKICETVAAQILIEHKLAKDEALKROSLGSLSPNSILAHGDIRNNKLNQLSVNLW HLAQRGSCADTRMYKTLIEYIKKOSKQPMTHLELALRLPLQTRT"
BASE COUNT 23581 a 19688 c 20974 g 21913 t
ORIGIN

Query Match 1.2%: Score 25; DB 9; Length 86156;
Best Local Similarity 100.0%: Pred No. 1.8;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 51552 AAAAAAAAAATAAATAAGAGAAAG 51576
1859 AAAAAAAAAATAAATAAGAGAAAG 1883
AAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

RESULT 5
AL360221 16616 bp DNA HTG 16-MAY-2001
LOCUS Homo sapiens chromosome 9 clone RP11-279E1, *** SEQUENCING IN
DEFINITION PROGRESS ***, 4 unordered pieces.
ACCESSION AL360221

VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL360221.11 GI:13157574
HTG: HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLPROP.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 166116)
Direct Submission
Submitted (15-MAY-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Feb 27, 2001 this sequence version replaced gi:13092300.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA279E1
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-termination Big Dye; 100% of reads
Consensus quality: 164733 bases at least Q40
Consensus quality: 165336 bases at least Q30
Consensus quality: 165634 bases at least Q20
Insert size: 165816; sum-of-ctrls
Insert size: 158936; 8.7% error; agarose-fp
Quality coverage: 6.79% in Q20 bases; sum-of-ctrls Quality.
Coverage: 7.13% in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 34340: contig of 34340 bp in length
* 34341 34440: gap of 100 bp
* 34441 100244: contig of 65804 bp in length
* 100245 100344: gap of 100 bp
* 100345 123592: contig of 23248 bp in length
* 123593 123692: gap of 100 bp
* 123693 166116: contig of 42424 bp in length.
Location/Qualifiers
1. 166116
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-279E1"
/clone_1lb="RPCT-11.1"
1. 34340
/note="assembly-fragment:02226
vector_side:left"
clone_end:SP6
vector_side:right"
34441. 100244
/note="assembly-fragment:01386
fragment_chain:1"
100345. 123592
/note="assembly-fragment:02487
fragment_chain:1"
123693. 166116
/note="assembly-fragment:01891
clone_end:AT7
vector_side:right"
BASE COUNT 50579 a 32344 c 31958 g 50934 t 301 others
ORIGIN
Query Match 1.2%: Score 25; DB 2; Length 166116;

Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1859 AAAAAAAAAATAAATAAAGAAAG 1883
|||||

Db 157781 AAAAAAAAAATAAATAAAGAAAG 157757

RESULT 6

AC013508 184666 bp DNA HTG 09-SEP-2000
Locus Homo sapiens clone RP11-11516, WORKING DRAFT SEQUENCE, 20 unordered
pieces.

AC013508 GI:10045243

AC013508.3 HTG: HTGS_PHAS1; HTGS_DRAFT.

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

COMMENT

TITLE

JOURNAL

Direct Submission
Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 9, 2000 this sequence version replaced gi:6479139.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www.seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Center project name: L1876

Center clone name: 115.1.6

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 159232 bases at least Q40

Consensus quality: 174889 bases at least Q30

Consensus quality: 180287 bases at least Q20

Insert size: 165000; agarose-efp

Insert size: 182766; sum-of-covs

Quality coverage: 5.9 in Q20 bases; sum-of-covs

Quality coverage: 5.3 in Q20 bases; sum-of-covs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will

FEATURES

source

* be preserved. 6241: contig of 6241 bp in length
1 6241: gap of 100 bp in length
6342 6341: gap of 100 bp in length
6342 6341: gap of 100 bp in length
7729 7828: gap of 100 bp in length
7829 8847: contig of 1019 bp in length
8848 8947: gap of 100 bp in length
8948 10136: contig of 1189 bp in length
10137 10236: gap of 100 bp in length
10237 11541: contig of 1305 bp in length
11542 11641: gap of 100 bp in length
11642 13230: contig of 1589 bp in length
13231 13330: gap of 100 bp in length
13331 14500: contig of 1170 bp in length
14501 14600: gap of 100 bp in length
14601 15869: contig of 1269 bp in length
15870 15969: gap of 100 bp in length
15970 17053: contig of 1084 bp in length
17054 17153: gap of 100 bp in length
17154 18324: contig of 1171 bp in length
18325 18424: gap of 100 bp in length
18425 19469: contig of 1045 bp in length
19470 19569: gap of 100 bp in length
19570 21310: contig of 1741 bp in length
21311 21410: gap of 100 bp in length
21411 22951: contig of 1541 bp in length
22952 23051: gap of 100 bp in length
23052 25121: contig of 2070 bp in length
25122 25221: gap of 100 bp in length
25222 27102: contig of 1881 bp in length
27103 27202: gap of 100 bp in length
27203 30450: contig of 3248 bp in length
30451 30550: gap of 100 bp in length
30551 129558: contig of 99008 bp in length
129559 129558: gap of 100 bp in length
129559 147795: contig of 1837 bp in length
147796 147895: gap of 100 bp in length
147896 175355: contig of 27460 bp in length
175356 175455: gap of 100 bp in length
175456 184666: contig of 9211 bp in length.

source

Location/Qualifiers
1.184666 "Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-11516"
/clone_lib="RPC1-11 Human Male BAC"
1.6241
/note="assembly-fragment
clone_end:SP6
vector_side:left"
6342.7728
/note="assembly-fragment"
7829.8847
/note="assembly-fragment"
8948.10136
/note="assembly-fragment"
10237.11541
/note="assembly-fragment"
11642.13230
/note="assembly-fragment"
13331.14500
/note="assembly-fragment"
14601.15869
/note="assembly-fragment"
15970.17053
/note="assembly-fragment"
17154.18324
/note="assembly-fragment"
18425.19469
/note="assembly-fragment"
19570.21310
/note="assembly-fragment"
21411.22951

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

```

* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved
  1. 39912: contig of 39912 bp in length
  39913 40012: gap of 100 bp
  40013 70805: contig of 30793 bp in length
  70806 70905: gap of 100 bp
  70906 109829: contig of 38924 bp in length
  109830 109929: gap of 100 bp
  109930 151328: contig of 41399 bp in length
  151329 151428: gap of 100 bp
  151429 166455: contig of 15027 bp in length
  166456 166555: gap of 100 bp
  166556 188120: contig of 21565 bp in length.
  Location/Qualifiers
    1..188120
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /chromosome="1"
      /clone="RP11-502P9"
      /clone_11b="RPCT-11.2"
      1..39912
        /note="assembly_fragment:02972
        fragment_chain:1"
        40013..70805
          /note="assembly_fragment:01087
          fragment_chain:1"
          70906..109829
            /note="assembly_fragment:01395
            fragment_chain:1"
            109930..151328
              /note="assembly_fragment:02833
              fragment_chain:2"
              151429..166455
                /note="assembly_fragment:00039
                fragment_chain:2"
                166556..188120
                  /note="assembly_fragment:01325
                  fragment_chain:2
                  clone_end:SP6
                  vector_side:right"
BASE COUNT      50383 a      39901 c      42215 g      55121 t      500 others
ORIGIN
Query Match      1.2%: Score 25; DB 2; Length 188120;
Best Local Similarity 100.0%: Pred. NO. 1.8;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY 1858 TAAAAAAATTAATTAATAAGAAAA 1882
|||||
Db 92166 TAAAAAAATTAATTAATAAGAAAA 92190
RESULT 8
AC093278 193305 bp DNA HTG 16-AUG-2001
LOCUS Homo sapiens chromosome 5 clone RP11-389C8, WORKING DRAFT SEQUENCE,
DEFINITION 33 unordered pieces.
ACCESSION AC093278
VERSION AC093278.1 GI:15193412
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEPIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 193305)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 5
Unpublished
2 (bases 1 to 193305)
DOE Joint Genome Institute.
REFERENCE
AUTHORS
TITLE
Direct Submission

```


variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated repeat sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: EM, EMBL; SW, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/Projects/Celegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr9> Rpl1-575C20 is from the library RPl1-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

FEATURES

SOURCE

1. 199517
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RPl1-575C20"
/clone_1lb="RPl1-11.2"
/clone_1lb="RPl1-11.2"
/note="Single clone region. Assembly confirmed by restriction digest data."

misc_feature

159357..159427
/note="Single clone region. Assembly confirmed by restriction digest data."

BASE COUNT 61043 a 40153 c 38840 g 59481 t
ORIGIN

Query Match 1.2%; Score 25; DB 9; Length 199517;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1859 AAAAAAATATAATTAAGCAAG 1883
|||||
Db 125079 AAAAAAATATAATTAAGCAAG 125103

RESULT 10
AL590624/C 211271 bp DNA HTG 02-MAY-2001
LOCUS Homo sapiens chromosome 1 clone RPl1-380B22, *** SEQUENCING IN
DEFINITION
ACCESSION AL590624 GI:13750986
VERSION AL590624.4
KEYWORDS HTG; HTGS_PHRASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 211271)
AUTHORS McLay, K.
TITLE Direct Submission
SUBMITTED (01-MAY-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
REQUESTS: clonerequest@sanger.ac.uk
On Apr 21, 2001 this sequence version replaced gi:13660991.

COMMENT

Center: Sanger Centre
Center code: SC
Web site: <http://www.sanger.ac.uk>

Contact: humquerry@sanger.ac.uk

----- Project Information
Center project name: BA380B22
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 21006; bases at least Q40
Consensus quality: 210412; bases at least Q30
Consensus quality: 210565; bases at least Q20
Insert size: 210871; sum-of-contigs
Insert size: 208059; 5.5% error; agarose-ff
Quality coverage: 8.15x in Q20 bases; sum-of-contigs Quality
coverage: 8.25x in Q20 bases; agarose-ff

----- NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 9007: contig of 9007 bp in length
* 9008 9107: gap of 100 bp
* 9108 60690: contig of 51583 bp in length
* 60691 60790: gap of 100 bp
* 60791 119563: contig of 58875 bp in length
* 119670 119769: gap of 100 bp
* 119770 169725: contig of 49956 bp in length
* 169726 169825: gap of 100 bp
* 169826 211271: contig of 41446 bp in length.

FEATURES

SOURCE

1. 211271
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RPl1-380B22"
/clone_1lb="RPl1-11.2"
/clone_1lb="RPl1-11.2"
/note="assembly-fragment:02974
clone_end:17
vector_side:left"

misc_feature

9108..60690
/note="assembly-fragment:01360
fragment_chain:1"

misc_feature

60791..119669
/note="assembly-fragment:03056
fragment_chain:1"

misc_feature

119770..169725
/note="assembly-fragment:01162
fragment_chain:1"

misc_feature

169826..211271
/note="assembly-fragment:00526
fragment_chain:1
clone_end:SP6
vector_side:right"

BASE COUNT 62583 a 47126 c 44740 g 56422 t 400 others
ORIGIN

Query Match 1.2%; Score 25; DB 2; Length 211271;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1858 TAAAAAATATAATTAAGCAAG 1882
|||||
Db 67484 TAAAAAATATAATTAAGCAAG 67460

RESULT 11

PGCBP2 1366 bp DNA INV 12-SEP-1997

```

REFERENCE 1 (bases 1 to 1366)
AUTHORS Ruedolph,B., Nolte,D. and Knapp,B.
TITLE Isolation of a third member of the Plasmodium falciparum
JOURNAL glycoprotein-binding protein gene family
MEDLINE 95198770
REFERENCE 2 (bases 1 to 1366)
AUTHORS Knapp,B.
TITLE Direct Submission
JOURNAL Submitted (22-JUN-1994) B. Knapp, Behringwerke AG, Research
Laboratories, P O Box 11 40, 3550 Marburg/Lahn, FRG

FEATURES
source
1..1366
/organism="Plasmodium falciparum"
/strain="FCBR (Colombia)"
/db_xref="taxon:5833"
/dev_strage="blood stage"
join(134..322,479..1306)
/gene="GBPH2"
join(134..322,479..1306)
/gene="GBPH2"
134..322
/gene="GBPH2"
/number=1
323..478
/number=1
479..1306
/gene="GBPH2"
/number=2

BASE COUNT 551 a 219 c 201 g 395 t
ORIGIN

Query Match 1.1%; Score 24; DB 3; Length 1366;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1859 AAAAAAAAAATAAATAAGAAAA 1882
Db 2 AAAAAAAAAATAAATAAGAAAA 25

RESULT 13
AC019982/c HTG 03-JAN-2006
DEFINITION AC019982 8077 bp DNA
AC019982 Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
AC019982 pieces.
AC019982 AC019982.1 GI:6664915
AC019982 HTG; HTGS_PHASE2.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 8077)
Adams,M. and Venter,J.C.
Direct Submission
Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
This sequence was identified as CDM.10211617 by the submitter.
For more information on this record e-mail to fly@celera.com.
NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
Location/Qualifiers
1..8077
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"

BASE COUNT 2477 a 1794 c 1761 g 2045 t
ORIGIN

```

[illegible]

SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus-1 (bases 1 to 114898)
AUTHORS Martindale,D.W., Wilson,M.D., Wang,D., Burke,R.D., Chen,X., Duronio,V. and Koop,B.F.
TITLE Comparative genome sequence analysis of the Williams syndrome region (LIMK1-RFC2) of human chromosome 7q11.23
JOURNAL Mamm. Genome 11 (10), 890-898 (2000)
MEDLINE 20458868
PUBMED 11003705
REFERENCE 2 (bases 1 to 114898)
AUTHORS Martindale,D.W., Wilson,M.D. and Koop,B.F.
TITLE Direct Submission
JOURNAL Submitted (01-APR-1999) Department of Biology, Centre for Environmental Health, University of Victoria, P.O. Box 3020, Victoria, BC V8W 3N5, Canada
FEATURES
source
1. 114898
/organism="Mus musculus"
/strain="ES-129/SV01"
/db_xref="taxon:10090"
/clone="Genome Systems Inc BAC17511"
join(2873..3130,4731..4827,18321..18459,18691..18800,20947..21153,22176..22281,22391..22557,25458..25641,26178..26264,27155..27286,29309..29368,29613..29678,31468..31624,33507..33562,33821..33978,34086..35418)
/gene="Limk1"
/product="Lim-kinase1"
2873..35418
/gene="Limk1"
join(3076..3130,4731..4827,18321..18459,18691..18800,20947..21153,22176..22281,22391..22557,25458..25641,26178..26264,27155..27286,29309..29368,29613..29678,31468..31624,33507..33562,33821..33978,34086..34248)
/gene="Limk1"
/codon_start=1
/product="Lim-kinase1"
/protein_id="AAD34858.1"
/db_xref="GI:4972949"
/translation="MRLLTCCITWREKREGESELPYCASGGRHYDGYIALNDWANDCRCCESYSIHOYEKDGOLFCKDKWAWYEGSCSHITKGLVWAGELKHPECFILACGNFTGDDDTITVHSHSLYCCGYVTVTPTVELLPSPSHIPIHTVLSIPASAGKRLSVSIDPFGPGCCTESHHTVRQGVDPGMSDVANSTHVGDRILENGTPIRNPVLEIDLIIQETSRLLQLEHDPDSLHGVSUPLSSVHTPSCQASASARQKPYRLSCSIDTSPGTSLSAPASQKDLGRSESLAVVCRRHLE RPSDLHGEVLGKGCPCGOAKVTHETGEVWAKELIFDEDTORTFEKVVACLEHPVILKEFIVLYDKRLNFTTEVYKGGTREGIIRKNDQYPMSCRVSPAKIASMAY LHSMTIIRHDINSHNCLVRENRYVADGLRIWIDENQSEDLRSIKKDRKRYT VGNPTFMAMENKRSYDEKDVPSFGVLCETIGRYNADPDYLPRTMDGLVNGF LDRICPNCPSPTFFITVRCDDLPFKRPSFKLEQWLETLMLHSLGHLPLGPOLEOL ERGFWETRYRGESSLPAHPEVPD"
/gene="Limk1"
35395..35400
join(52154..52241,63700..63887,65870..65934,66013..66109,66975..67034,69389..69526,69827..71599)
/gene="Wbscr1"
/product="Wbscr1"
join(52154..52241,63700..63887,65870..65934,66013..66109,69389..69526,69827..71599)
/gene="Wbscr1"
/product="Wbscr1 alternative spliced product"
52154..71599
/gene="Wbscr1"
join(52183..52241,63700..63887,65870..65934,66013..66109,69389..69526,69827..69966)
/gene="Wbscr1"
/codon_start=1
/product="Wbscr1 alternative spliced product"
/protein_id="AAD34860.1"
/db_xref="GI:4972951"
/translation="MADFITYDDRAVSSFGGGRSGSAGHSRQKRLPTPEPYTA

CDS
YVGNLFNTVGGIDDAIFKDLISRSVRLVBDKDTDKFKGCFYEPDEVDLSKALTYD
CALLGDRSLRYDIAEGRKQKGGCFKGGPDGDDLDLGGGSGSRPDGAPPM
GSRDDGPPPLGSMWDRPEEERARPRQLKPRVATPLNQVANPNSAIFGARGP
REEVQKQKQ
join(52183..52241,63700..63887,65870..65934,66013..66109,66975..67034,69389..69526,69827..69966)
/gene="Wbscr1"
/codon_start=1
/product="Wbscr1"
/protein_id="AAD34859.1"
/db_xref="GI:4972950"
/translation="MADFITYDDRAVSSFGGGRSGSAGHSRQKRLPTPEPYTA
YVGNLFNTVGGIDDAIFKDLISRSVRLVBDKDTDKFKGCFYEPDEVDLSKALTYD
CALLGDRSLRYDIAEGRKQKGGCFKGGPDGDDLDLGGGSGSRPDGAPPM
GSRDDGPPPLGSMWDRPEEERARPRQLKPRVATPLNQVANPNSAIFGARGP
REEVQKQKQ
71576..71581
/gene="Wbscr1"
join(76469..76586,82545..82667,83281..83320,84654..84697,84855..84915,85083..85114,85513..85543,86674..86729,87995..88054,88305..88376,88806..88914,90506..91457)
/gene="Wbscr5"
/product="Wbscr5 alternative spliced product"
join(76469..76586,82545..82667,83281..83320,84654..84697,84855..84915,85083..85114,85594..85429,85513..85543,86674..86729,87995..88054,88305..88376,88806..88914,90506..91457)
/gene="Wbscr5"
/product="Wbscr5"
76469..91457
/gene="Wbscr5"
join(82571..82667,83281..83320,84654..84697,84855..85083,85114,85513..85543,86674..86729,87995..88054,88305..88376,88806..88888)
/gene="Wbscr5"
/codon_start=1
/product="Wbscr5 alternative spliced product"
/protein_id="AAF75559.1"
/db_xref="GI:8470891"
/translation="MSAELELMPYSGILLILGATAMLCVHCSPGVKRNKITYEOR
NQENQSSAAQOTYSLARQVPGQMTAPKSPERKNKLEFHLGSDNEDPAAY
DIPITNYVMGQFQKPSDDDSNSTEYLVKPSFTPEGVNDVFDYQNSYIHQWRS
KRTMGAPKLSGSPDEPDYVNGDAALNT"
join(82571..82667,83281..83320,84654..84697,84855..84915,85083..85114,85394..85429,85513..85543,86674..86729,87995..88054,88305..88376,88806..88888)
/gene="Wbscr5"
/codon_start=1
/product="Wbscr5"
/protein_id="AAF75558.1"
/db_xref="GI:8470890"
/translation="MSAELELMPYSGILLILGATAMLCVHCSPGVKRNKITYEOR
NQENQSSAAQOTYSLARQVPGQMTAPKSPERKNKLEFHLGSDNEDPAAY
DIPITNYVMGQFQKPSDDDSNSTEYLVKPSFTPEGVNDVFDYQNSYIHQWRS
KRTMGAPKLSGSPDEPDYVNGDAALNT"
ONSVS:HQWRESKRTVGAPWSLSGSPDEPDYVNGDAALNT"
complement(join(93152..93630,94484..94597,96116..96196,97205..97270,98133..98290,100260..100360,100502..100603,102020..102126,102508..102549,104514..104583,108595..108710))
/gene="Rfc2"
/product="replication factor C, 40kDa subunit"
complement(93152..108710)
/gene="Rfc2"
complement(93165..93169)
/gene="Rfc2"
complement(join(93520..93630,94484..94597,96116..96196,97205..97270,98133..98290,100260..100360,100502..100603,102020..102126,102508..102549,104514..104583,108595..108692))
/gene="Rfc2"
/codon_start=1
/product="replication factor C, 40kDa subunit"
/protein_id="AAD34861.1"

/db.xref="GI:4972952"
/translation="MEVQSGCDPSESGAOPSPVPSKTAGHYELPWVEKYRPLKNE
IVSNEPTVSRLEVFPAREGNVPNIITAGPCTGKTSTICLARALIGPALKDAVELNA
SNDRGIDVVRNKIKMEAQKVTLPKGRHRTIILDEADSTDGAQALRTMETISKTT
RFALACNASDKLIEPIQSRCAVLRKTLDAQVLRMLNVIKEKVPYTDGLEALIF
TAQGDNRQALNNLQSTFSGFGYINSENVKVCDEPHPLLVEMIQHCVDANIDEAYRI
LAHLMHIGYSPEDVIGNIFRYCKTFPMAEYLKLEFIKEIGYTHMKVAGVNSLLQMG
LIALRCQKTMAPVAS"
BASE COUNT 28567 a 28771 c 28573 g 28987 t
ORIGIN

Query Match 1.1%; Score 24; DB 10; Length 114898;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1861 AAAAAATTAATAAAGAAAAAGA 1884
Db 33257 AAAAAATTAATAAAGAAAAAGA 33280

Search completed: April 16, 2002, 18:18:06
Job time: 6547 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2002, 16:46:45 ; Search time 120.58 Seconds

(without alignments)
14859.925 Million cell updates/sec

Title: US-09-636-826-1

Perfect score: 1 GATATCACACATTTCGTCCAT.....TCGGCTTCTGGGCTCTCT 2090

Sequence: OLIGO_NUC

Gapop 60.0 , Capext 60.0

Scoring table: 930621 segs, 428662619 residues

Searched:

Word size : 0

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_1101.*

1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.*

* Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2090	100.0	2090	22	AA661197
2	23	1.1	162450	21	AA286967
3	22	1.1	543	22	AAH12273
4	22	1.1	1562	22	AA632540
5	22	1.1	3021	22	AA632539
6	22	1.1	6542	22	AA632538
7	22	1.1	50000	21	AA664140
8	21	1.0	183	22	AA616159
9	21	1.0	885	22	AAH04005
10	21	1.0	1381	15	AAQ77883
11	21	1.0	1381	17	AA277655

12	21	1.0	1418	17	AA27738	Neural thread prot
13	21	1.0	1423	21	AA297059	Human secreted pro
14	21	1.0	1682	21	AA618279	Lung cancer associ
15	21	1.0	1737	21	AA256739	Human transmembran
16	21	1.0	513445	22	AA616173	Soybean 318013 reg
17	20	1.0	182	22	AA616174	Soybean 515002 reg
18	20	1.0	197	21	AA605247	Human secreted pro
19	20	1.0	216	21	AA606620	Human colon cancer
20	20	1.0	319	22	AA688025	Peppermint plant o
21	20	1.0	332	21	AA619809	Human secreted pro
22	20	1.0	404	18	AA69381	Murine metastatic
23	20	1.0	445	21	AA615573	Human prostate can
24	20	1.0	474	22	AA638578	Probe #7264 used t
25	20	1.0	568	21	AA68015	568 bp Candida alb
26	20	1.0	598	21	AA67965	598 bp Candida alb
27	20	1.0	633	13	AA620241	Pis6 gene. Plasm
28	20	1.0	781	20	AA615116	Human gene express
29	20	1.0	946	21	AA658016	946 bp Candida alb
30	20	1.0	2733	19	AA613832	Homo sapiens telom
31	20	1.0	2733	19	AA605369	Human telomerase p
32	20	1.0	2809	22	AA614068	Human cDNA sequenc
33	20	1.0	4233	21	AA629018	Rat mdrlb2 (multis
34	20	1.0	4233	22	AA627498	Rat mdrlb2 (multis
35	20	1.0	9789	17	AA611852	cDNA encoding plas
36	20	1.0	127197	22	AA61370	Soybean 515002 reg
37	19	0.9	47	21	AA688270	Human map-related
38	19	0.9	62	21	AA64614	Au rich sequence i
39	19	0.9	212	21	AA614970	Human secreted pro
40	19	0.9	245	21	AA614835	Human secreted pro
41	19	0.9	292	21	AA611355	Aspergillus niger
42	19	0.9	312	22	AA611673	Human cDNA clone (
43	19	0.9	313	21	AA643641	Human secreted exp
44	19	0.9	363	21	AA613310	Human secreted pro
45	19	0.9	403	22	AA629465	Drosophila melano

ALIGNMENTS

RESULT 1	
ID	AA661197 standard; DNA; 2090 BP.
XX	
AC	AA661197;
XX	
DT	23-MAY-2001 (first entry)
XX	
DE	B. vulgaris V-ATPase subunit C isoform 2 promoter DNA.
XX	
KW	Promoter; V-ATPase; subunit C; isoform 2; plant; transgenic plant;
KW	V-type adenosine triphosphatase; stress resistance; insect resistance;
KW	hericide resistance; disease resistance; salt stress; constitutive;
KW	nutrient deficiency; injury; infection; selection marker; ss.
XX	
OS	Beta vulgaris.
XX	
PN	WO200114572-A2.
XX	
PD	01-MAR-2001.
XX	
PF	10-AUG-2000; 2000MO-EP07774.
XX	
PR	26-AUG-1999; 99DE-1040648.
PR	26-AUG-1999; 9905-0150887.
XX	
PA	(BADI) BASF AG.
XX	
PI	Duvenig E, Rausch T;
XX	
DR	WPI; 2001-218455/22.
XX	
PT	New DNA construct containing a plant V-type adenosine triphosphatase
PT	promoter, useful for recombinant protein expression or producing

PT stress-resistant plants
XX
PS Claim 6; Page 65-66; 69pp; German.
XX

This invention describes a novel DNA construct (I) which comprises a plant V-ATPase (V-type adenosine triphosphatase) promoter (P), or its functional equivalent, coupled to a heterologous gene. The invention also describes (1) a polynucleotide (II) containing the sequence of (P) from the Beta vulgaris V-ATPase subunit c isoform 2 (2090 bp sequence of (1), reproduced), or its functional equivalents; (2) a recombinant vector containing (I); (3) a microorganism transformed with the vector of (2); (4) a transgenic plant cell or protoplast that contains (I); (5) a transgenic plant containing (I) in its genome; (6) controlling (M1) expression of (II) in a plant cell or protoplast by transforming with (I) then exposing to an (a)biotic stress that controls expression of (II); (7) method (M2) similar to (M1), but applied to regenerated transgenic plants; (8) preparation (M3) of recombinant proteins (IV) by methods (M1) and (M2); (9) (II) produced by method (M3); and (10) plant cells or protoplasts, and regenerated plants, transformed with (I) and resistant to (a)biotic stresses. (I) and (II) are used to direct expression of recombinant proteins in plants, or their cells or protoplasts. (II) is used to provide expression of genes in plants under (a)biotic stresses, such as genes that impart resistance to insects, herbicides, diseases, or selection genes. (IV) is used to produce plants with increased resistance to stress (especially salt but also nutrient deficiency, injury, infection, presence of herbicides). P are very strong, constitutive promoters, effective in all parts of a plant, and they become more active, or are not repressed, under conditions of stress, so are better than known plant or viral promoters for expressing selection markers and resistance genes.

Sequence 2090 BP; 794 A; 354 C; 271 G; 671 T; 0 other;

Query Match 100.0%; Score 2090; DB 22; Length 2090;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 2090; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATATCACACATTCGTCATCGAGCATTTGGCACTTCAATAGGTACGTAATCTTTT 60
Db 1 gatacacacatctgcacatcgacgatttgcgaacttcaaatagtagtaattctttt 60
QY 61 AATCTTCAAAATTTATTCACATTTAGTACTATTATGTATGAAAAAAGGAGTAATA 120
Db 61 aatcttcaaaatltatcacatgacgactatgtagtgaaaaaagagataaata 120
QY 121 GGAATTAGTTAAAAAGGCTTTTATATATATAGACTTAATTTGATTCATTTTCAATAT 180
Db 121 ggaattagttgaaaaagggtttatataataatgacttaatttgatcatttccatatt 180
QY 181 CTGAAAACAGGATATGATGAAATTTGATTCATTTATGACACTGATGAAAAAATTACGA 240
Db 181 ctgaaaacaggatgtagtgaatttgaatttcaatgacactgataaaaaaattaca 240
QY 241 TTATAGTCTTTTAAAAATTCATATATAATTTTGGCCAAAACCTTTGGCAAAATATC 300
Db 241 ttatagctttttttaaattccaataataattttggccaaaaccttttgcaaaatattc 300
QY 301 CATGTTGGGAATTAATTTTGA AAAACAATATCAAACTTTTGGCAACAATT 360
Db 301 catgttcggaataaatttgaaaaaacaatacaaacctttttgcaaaacctttt 360
QY 361 TCAAAAATCCATTTTGA AAAAAAATTTACATTTAGCTTGGCAATCAATTTGGTAT 420
Db 361 tcaaaaatccatttgcgaaaaaaaatttaacttaacttgcgaaatcaaatgtgta 420
QY 421 GAAAAATTTAAATTTCTTTTACCTTATATTAAGTCAAAAGTGTAAATTTAGAAAA 480
Db 421 gaaaaatttaaatcttcccttcaactataatgaaactcaaggtttaaattttagaaa 480
QY 481 GGAGAAAAATTAATAATGACATTTGATGCAAAATGATGATGAAAAAATTTTAA 540
Db 481 ggagaaaaataataatgacatttcaatgcgaaatcaaatgtgtatgaaaaaattcaaa 540

QY 541 TTTTATTTTAAATATTAATGAAATTCAAAGCTTTAAAAATTTAGAAAAAGAGAAAAATTA 600
Db 541 ttttattttaataataatgaaattcaaaagtgttaaatttagaaagagaaaaattaa 600
QY 601 AATGACATTTTCAATTTCAAAATCGATTTGTATGAAAATTTTAAATTTTATTTCAATA 660
Db 601 aatgacatttcaatcctcaaaatcgatctgtatgaaaatcaaatcttattcaata 660
QY 661 TAAATGAACTCAAGTGTACATTTAGAAAAAGGAAAAATTTAAATGATGAAAAATT 720
Db 661 taattgaaactcaaggtgtacattttagaaaagagaaaatttaaatgataaat 720
QY 721 GTAAATATCAATTTTGA AAAATCAGAAATTTAGAGTTATACAGAGAAAAAATCGAATT 780
Db 721 gtaaaatcaaatlttgaaatcagaattttagaagtttagaagaaaaaaacggaatt 780
QY 781 GTCTTATACCTTTTGGTACAAATTTTGGATCAATAAGAAATTTACTGAAATCGATATCA 840
Db 781 gtcttataacttttgggtacattttgggattcataagaatttactgaatccatata 840
QY 841 AAATATTTTAAATTAATAAATGATTAACCAAAAAAGAAAGAAACATGAGATATTC 900
Db 841 aaactatataaatttacaataatgaaataaaacaaaaaagaacacgacgatacttc 900
QY 901 GTAAAGAACATCACTACTGATTAATAAAGAACATGCCATTTAGATTGAGAAACAAAA 960
Db 901 gtaagaacatcatactgattataaagaacatcgcatatttagaattgagaacaaaa 960
QY 961 ACTTTTCAAAATCCAAAAATGATTAACAAATCAACAAAGAAATGAAATCTTATCA 1020
Db 961 actttaaaaaatcacaaaaatgatacaacaacaaagaacatgaaataacttattca 1020
QY 1021 CAAATGAGGTGACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
Db 1021 caaatgaggtgacttaataataataataataataataataataataataataata 1080
QY 1081 AGGCTTAAAAATATCATCGACAGCATCGCTGCACAAAAAGCTAGTGTATATACAA 1140
Db 1081 aggccttaaaaatatacatcgacagcatcgctgcacaaaagactaggttaatacaaa 1140
QY 1141 GTCAACAACATCACTACTGATTTTCAATTTAGTTCATTTGATGAGGCGTAAACGTTTTC 1200
Db 1141 gtcaacaacatcaactgatttcaattttagtgcctccatttggtagggcgtaaacgttttc 1200
QY 1201 CCGAACAATATTTTCTCATTTTATTTAGTTTATCATTTGTTGTTGACAAAAAGATGTA 1260
Db 1201 ccggaacaatatttctcatatttttagtttcaatgtttgttgacaaaagatgttaa 1260
QY 1261 AACGTTTTCCTTGGGGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1320
Db 1261 aacgtttcccttggggtaaatattactcttccaatgatacgaataacatttccattcaa 1320
QY 1321 AATGAAGGAAATGTTTCTTATATCTGTTTGTACATTTTCTACCTTCTTAC 1380
Db 1321 aatgaaggaaatggtttcttattatctgtttgttacaatttccactactcccttacc 1380
QY 1381 TTCCCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1440
Db 1381 ttcccttttatttatttatttatttatttatttatttatttatttatttatttatttattt 1440
QY 1441 ACTAATTTTGAATGTGTTCATTTTCAATTTTGAATTTTGAATTTTGAATTTTGAATTT 1500
Db 1441 actaattttgaaattgttcttcatatttgaatttgaatttgaatttgaatttgaatttgaatt 1500
QY 1501 GAAATGTTTATACCTTACCAACAGAGCTTCGTTGCAATGATGATGATGATGATGATG 1560
Db 1501 gaaatgttttacccccaacagagcttcggttgccatgaaatgaaatgaaatgaaatgaaat 1560
QY 1561 CAAATTCGAAATGATGTTTACGTTTCTTCAATTTGTTTGAATTTGATGATGATGATGAT 1620
Db 1561 caaattcgaaatgattgttttaccgtttcccaatgtttcgaattgattcacaacgaatttc 1620

CC	genomic sequence of the invention. The RBP-7 coding sequence and	
PT	differentiation -	
XX		
PS	Claim 1, Page 118-163; 223pp. English.	
XX		
CC		
XX		
PT	Novel nucleic acid and polymorphic markers used for diagnosis of	
XX	diseases, especially those involving abnormal cell proliferation and	
XX		
DR	WPI; 2000-117170/10.	
XX		
PI	Bougueleret L;	
XX		
PA	(GEST) GENSET.	
XX		
PR	10-DEC-1998; 9805-0111909.	
XX		
XX	30-JUN-1998; 9805-0091315.	
XX		
PF	30-JUN-1999; 99WO-IB01242.	
XX		
PD	06-JAN-2000.	
XX		
PN	WO200000607-A1.	
XX		
OS	Homo sapiens.	
XX		
XX		
DE	Retinoblastoma binding protein-7 genomic DNA sequence.	
XX		
DT	16-MAY-2000 (first entry)	
XX		
XX		
KW	RBP-7; retinoblastoma binding protein-7; abnormal cell proliferation;	
KW	diagnosis; therapy; cell differentiation; thyroid hyperplasia; psoriasis;	
KW	benign prostate hypertrophy; cancer; sarcoma; neoplasm; leukaemia;	
KW	lymphoma; ds.	
XX		

Query Match	1.1%	Score 23	DB 21	Length 162450
Best Local Similarity	100.0%	Pred. No. 0.99		
Matches	23	Conservative	0	Mismatches
			0	Indels
			0	Gaps
QY	1858	TAATAAAATTAATTAATGA	1880	
Db	137096	taataaataataataaaga	137118	

RESULT 3
 ID AAH12273
 XX AAH12273 standard; cDNA; 543 bp.
 AC
 XX AAH12273;
 DT
 XX 26-JUN-2001 (first entry)
 XX
 DE Human cDNA clone (3'-primer) SEQ ID NO:9108.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 CS Homo sapiens.
 XX
 FN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 94JP-0248036.
 PR 27-AUG-1999; 94JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0189767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI: 2001-318749/34.
 XX
 PR Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 TT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 PS
 Claim 3; SEQ ID 9108; 2537pp + CD ROM; English.
 CC
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AA92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX Sequence 543 BP; 176 A; 115 C; 70 G; 176 T; 6 other;

Query Match 1.1%; Score 22; DB 22; Length 543;

Best Local Similarity 100.0%; Pred. No. 3.4; Mismatches 0; Indels 0; Gaps 0;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 420 TGAATAATTTAAATTCCTT 441
 DB 511 tgaataatttaatttcctt 532

RESULT 4

AA92540/C
 ID AAF32540 standard; DNA; 1562 BP.

XX AAF32540;

DE 19-APR-2001 (first entry)

XX Mouse P/Q-type calcium channel upstream sequence SEQ ID NO:3.

KW Mouse; P/Q-type calcium channel upstream sequence; brain; neuron;
 KM P/Q-type calcium channel a1A; cerebral neuron; ds.

XX Mus sp.

PN JP2000316583-A.

PD 21-NOV-2000.

PF 14-MAY-1999; 99JP-0133421.

PR 14-MAY-1999; 99JP-0133421.

PA (EISA) EISAI CO LTD.

DR WPI; 2001-066178/08.

XX A vector expressing specifically in brain

PS Claim 3; Page 12; 23pp; Japanese.

CC The present sequence represents a mouse P/Q-type calcium (Ca) channel
 CC upstream sequence. The present invention describes a method which can
 CC be used for expressing a heterologous gene in neurons especially
 CC cerebral neurons.

XX Sequence 1562 BP; 274 A; 482 C; 412 G; 394 T; 0 other;

Query Match 1.1%; Score 22; DB 22; Length 1562;

Best Local Similarity 100.0%; Pred. No. 3.3; Mismatches 0; Indels 0; Gaps 0;

QY 1863 AAAATAAATAAAGAAAGA 1884
 ||||||||||||||||||

DB 1213 AAAATAAATAAAGAAAGA 1192

RESULT 5

AA92539/C
 ID AAF32539 standard; DNA; 3021 BP.

XX AAF32539;

DE 19-APR-2001 (first entry)

XX Mouse P/Q-type calcium channel upstream sequence SEQ ID NO:2.

KW Mouse; P/Q-type calcium channel upstream sequence; brain; neuron;
 KM P/Q-type calcium channel a1A; cerebral neuron; ds.

XX Mus sp.

PN JP2000316583-A.

PD 21-NOV-2000.

PF 14-MAY-1999; 99JP-0133421.

PR 14-MAY-1999; 99JP-0133421.

PA (EISA) EISAI CO LTD.

DR WPI; 2001-066178/08.

XX A vector expressing specifically in brain
 PS Claim 2; Page 10-11; 23pp; Japanese.

CC The present sequence represents a mouse P/Q-type calcium (Ca) channel
 CC upstream sequence. The present invention describes a method which can
 CC be used for expressing a heterologous gene in neurons especially
 CC cerebral neurons.

XX Sequence 3021 BP; 693 A; 810 C; 707 G; 811 T; 0 other;

Query Match 1.1%; Score 22; DB 22; Length 3021;

Best Local Similarity 100.0%; Pred. No. 3.2; Mismatches 0; Indels 0; Gaps 0;

QY 1863 AAAATAAATAAAGAAAGA 1884
 DB 2672 AAAATAAATAAAGAAAGA 2651
 ||||||||||||||||||

RESULT 6

AA92538/C
 ID AAF32538 standard; DNA; 6542 BP.

XX AAF32538;

DE 19-APR-2001 (first entry)

XX Mouse P/Q-type calcium channel upstream sequence SEQ ID NO:1.

KW Mouse; P/Q-type calcium channel upstream sequence; brain; neuron;
 KM P/Q-type calcium channel a1A; cerebral neuron; ds.

XX Mus sp.

PN JP2000316583-A.

PD 21-NOV-2000.

PF 14-MAY-1999; 99JP-0133421.

PR 14-MAY-1999; 99JP-0133421.

```

XX (EISA ) EISA CO LTD.
XX
XX WPI: 2001-066178/08.
XX
XX A vector expressing specifically in brain
XX
XX Claim 1; Page 8-10; 23pp; Japanese.
XX
XX The present sequence represents a mouse P/Q-type calcium (Ca) channel
XX upstream sequence. The present invention describes a method which can
XX be used for expressing a heterologous gene in neurons especially
XX cerebral neurons.
XX
XX Sequence 6542 BP; 1642 A; 1688 C; 1567 G; 1645 T; 0 other;

Query Match
Best Local Similarity 100.0%; Score 22; DB 22; Length 6542;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1863 AAAAATAAATAAAGAGAGAGA 1884
   |||||||||||||||||||
DB 5925 AAAAATAAATAAAGAGAGAGA 5904

RESULT 7
AAA64140
ID AAA64140 standard; DNA: 50000 BP.
XX
XX AAA64140;
XX
XX 20-DEC-2000 (first entry)
XX
XX Nucleotide sequence of a beta-tubulin antigen.
XX
XX Beta-tubulin antigen; inner ear protein; Meniere's disease; autoantibody;
XX chronic ear disease; autoimmune disease; ss.
XX
XX Homo sapiens.
XX
XX WO200050593-A1.
XX
XX 31-AUG-2000.
XX
XX 25-FEB-2000; 2000WO-US04795.
XX
XX 25-FEB-1999; 99US-0121549.
XX
XX (UYTE-) UNIV TENNESSEE RES CORP.
XX
XX YOO TJ;
XX
XX WPI: 2000-558400/51.
XX
XX New beta-tubulin antigen in the membranous structure of the inner ear,
XX reactive with antibodies of patients with Meniere's disease, for
XX diagnosing Meniere's disease and distinguishing this disease from other
XX autoimmune ear diseases
XX
XX Claim 3; Page 74-97; 115pp; English.
XX
XX The present sequence encodes a beta-tubulin antigen. The protein is
XX an antigen of the membranous structure of the inner ear protein, and
XX is reactive with antibodies from patients having Meniere's disease.
XX Meniere's disease is a chronic ear disease with unknown etiology.
XX Serum from patients suffering from this disease contain autoantibodies
XX against a 30 kDa cochlear protein antigen. The disease is believed to be
XX an autoimmune disease. The beta-tubulin antigen is useful as a target
XX substance in diagnosing or detecting Meniere's disease and in
XX distinguishing this disease from other autoimmune ear diseases.
XX
XX Sequence 50000 BP; 17281 A; 9480 C; 8791 G; 14448 T; 0 other;

```

```

Query Match
Best Local Similarity 100.0%; Score 22; DB 21; Length 50000;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1859 AAAAATAAATAAAGAGAGA 1880
   |||||||||||||||||||
DB 45486 aaaaaataataataaagaa 45507

RESULT 8
AAI61659
ID AAI61659 standard; DNA: 183 BP.
XX
XX AAI61659;
XX
XX 16-OCT-2001 (first entry)
XX
XX Soybean 318013 region A3 DNA, SEQ ID NO: 290.
XX
XX Soybean; antihelminthic; gene therapy; soybean cyst nematode; SCN;
XX SCN resistance; Rhg1; Rhg4; SCN resistant allele; plant breeding;
XX 240017 region G3; 318013 region A3; 515002 region G2; ds.
XX
XX Glycine max.
XX
XX WO200151627-A2.
XX
XX 19-JUL-2001.
XX
XX 05-JAN-2001; 2001WO-US00552.
XX
XX 07-JAN-2000; 2000US-0174880.
XX
XX (MONS ) MONSANTO CO.
XX
XX Hauge BM, Wang ML, Parsons JD, Parnell LD;
XX
XX WPI: 2001-425872/45.
XX
XX New purified nucleic acid for producing a soybean plant having soybean
XX cyst nematode resistance and for use in plant breeding programs
XX
XX Claim 25; Page 1103; 1353pp; English.
XX
XX The invention relates to nucleic acid molecules from regions of the
XX soybean genome which are associated with soybean cyst nematode (SCN)
XX resistance. The nucleic acids are used to transform plants, and can
XX produce soybean plants having an rhg1 or an Rhg4 SCN resistant allele.
XX The nucleic acids can be used for investigating rhg1 or Rhg4 haplotypes
XX of soybean plants and for introgressing SCN resistance or partial SCN
XX resistance into soybean plants. They can also be used in plant breeding
XX programmes. The invention also relates to proteins encoded by such
XX nucleic acid molecules, as well as antibodies capable of recognising
XX these proteins. The present sequence is a nucleic acid molecule
XX provided in the specification.
XX
XX Sequence 183 BP; 74 A; 12 C; 20 G; 77 T; 0 other;

Query Match
Best Local Similarity 100.0%; Score 21; DB 22; Length 183;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1862 AAAAATAAATAAAGAGAGA 1882
   |||||||||||||||||||
DB 84 aaaaaataataataaagaaa 104

RESULT 9
AAH04005
ID AAH04005 standard; CDNA; 885 BP.

```

```

XX AA04005;
AC
XX 26-JUN-2001 (first entry)
DT
XX Human CDNA clone (5'-primer) SEQ ID NO:840.
DE
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
KM
XX Homo sapiens.
OS
XX EPI074617-A2.
FN
XX 07-FEB-2001.
PD
XX 28-JUL-2000; 2000EP-0116126.
PF
XX 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
PA
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
PI
XX WPI; 2001-318749/34.
DR
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX Claim 1; SEQ ID 840; 2537PP + CD ROM; English.
PS
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AA03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human CDNA sequences; AA032446 to
CC AA035893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 885 BP; 307 A; 145 C; 200 G; 229 T; 4 other;
SQ

```

```

AA077883/c
ID AA077883 standard; CDNA; 1381 BP.
XX
XX AA077883;
AC
XX 06-JUL-1995 (first entry)
DT
XX
XX Neural thread protein AD10-7 CDNA.
DE
XX
XX Neural thread protein AD10-7; Alzheimer's; neuroectodermal tumours;
KM malignant astrocytomas; glioblastomas; ss.
XX
XX Rattus rattus.
OS
XX WO9423756-A.
FN
XX 27-OCT-1994.
PD
XX 20-APR-1994; 94WO-US04321.
PF
XX 20-APR-1993; 93US-0050559.
PR
XX (GEHO ) GEN HOSPITAL CORP.
PA
XX De LA MONTE SM, Wands JR;
PI
XX WPI; 1994-341497/42.
DR
XX
XX Detection of neural thread proteins - to detect sporadic and
PT familial Alzheimer's disease, neuroectodermal tumours, malignant
PT astrocytomas and glioblastomas (Eng).
PT
XX
XX Example 4; Fig 16R; 158pp; English.
PS
XX
XX AA077883 is the AD10-7 neural thread protein (NTP) cDNA. This
CC sequence was used in the development of an antibody dependent
CC method, for the detection of NTPs. This new method could be
CC used to diagnose Alzheimer's disease (differentiating between
CC sporadic and familial), neuroectodermal tumours, malignant
CC astrocytomas and glioblastomas.
CC
XX
XX Sequence 1381 BP; 296 A; 385 C; 301 G; 399 T; 0 other;
SQ

```

```

Query Match 1.0%; Score 21; DB 15; Length 1381;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1858 TAAATAAATAAATAAATAAAG 1878
DB 450 TAAATAAATAAATAAATAAAG 430

```

```

RESULT 11
AAT27765/c
ID AAT27765 standard; CDNA; 1381 BP.
AC
XX AAT27765;
XX
XX 14-NOV-1996 (first entry)
DT
XX
XX AD 10-7 human neural thread protein clone (partial sequence).
DE
XX
XX Neural thread protein; NTP; diagnosis; detection;
KM Alzheimer's disease; neuroectodermal tumour; malignant astrocytoma;
KM monoclonal antibody; binding fragment; ds.
XX
XX Homo sapiens.
OS
XX WO9615272-A1.
FN
XX 23-MAY-1996.
PD
XX
XX

```

RESULT 10

```

PF 14-NOV-1995; 95WO-US17111.
XX
XX 14-NOV-1994; 94US-0340426.
XX
XX (GEHO ) GEN HOSPITAL CORP.
PA
XX De LA MONTE S, Wands JR;
XX
XX WPI; 1996-259865/26.
DR
XX
XX Detection of neural thread protein in diagnosis of Alzheimer's
PT disease - also NTP DNA and protein sequences used in gene and
PT anti:sense therapy
XX
XX Example 4c; Figure 16R; 238pp; English.
PS
XX A method for detecting the presence of neural thread protein (NTP)
CC having a molecular weight of 8, 14, 17, 21, 26 or 42 kD in a human
CC subject comprises (a) contacting a sample from a human subject that
CC is suspected of containing the NTP with at least one molecule
CC capable of binding to the protein; and (b) detecting any of the
CC molecule bound to the protein. The binding molecule is selected
CC from an antibody free of natural impurities, a monoclonal antibody
CC or a binding fragment of either of these. The method may be used for
CC diagnosing the presence of Alzheimer's disease, neuroectodermal
CC tumours and a malignant astrocytoma in a human. A number of clones
CC of neural thread protein were isolated from healthy 17-18 week old
CC foetal human brain (HB) 2 year old temporal lobe neocortex and end
CC stage Alzheimer's disease (AD) cerebral cortex. See AAT27753-75.
XX
SQ Sequence 1381 BP; 295 A; 385 C; 301 G; 400 T; 0 other;

Query Match 1.0%; Score 21; DB 17; Length 1381;
Best Local Similarity 100.0%; Pred.No.9.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1858 TAAAAAAATAAATAAAG 1878
DB 450 TAAAAAAATAAATAAAG 430

RESULT 12
AAT27738/c
ID AAT27738 standard; DNA; 1418 BP.
XX
XX AAT27738;
AC
XX 15-NOV-1996 (first entry)
DT
XX
XX Neural thread protein coding sequence.
DE
XX
XX Neural thread protein; NTP; diagnosis; detection;
KM Alzheimer's disease; neuroectodermal tumour; malignant astrocytoma;
KM monoclonal antibody; binding fragment; ds.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX 14..1207
XX CDS /*tag= a
XX /product= Neural thread protein.
XX
XX W09615272-A1.
XX
XX 23-MAY-1996.
XX
XX 14-NOV-1995; 95WO-US17111.
XX
XX 14-NOV-1994; 94US-0340426.
XX
XX (GEHO ) GEN HOSPITAL CORP.
XX

```

```

PI De LA MONTE S, Wands JR;
XX
XX WPI; 1996-259865/26.
XX
XX P-PSDB; AAR95913.
DR
XX
XX Detection of neural thread protein in diagnosis of Alzheimer's
PT disease - also NTP DNA and protein sequences used in gene and
PT anti:sense therapy
XX
XX Claim 24; Page 168-170; 238pp; English.
PS
XX
XX A method for detecting the presence of neural thread protein (NTP)
CC having a molecular weight of 8, 14, 17, 21, 26 or 42 kD in a human
CC subject comprises (a) contacting a sample from a human subject that
CC is suspected of containing the NTP with at least one molecule
CC capable of binding to the protein; and (b) detecting any of the
CC molecule bound to the protein. The binding molecule is selected
CC from an antibody free of natural impurities, a monoclonal antibody
CC or a binding fragment of either of these. The method may be used for
CC diagnosing the presence of Alzheimer's disease, neuroectodermal
CC tumours and a malignant astrocytoma in a human.
XX
SQ Sequence 1418 BP; 302 A; 396 C; 315 G; 405 T; 0 other;

Query Match 1.0%; Score 21; DB 17; Length 1418;
Best Local Similarity 100.0%; Pred.No.9.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1858 TAAAAAAATAAATAAAG 1878
DB 445 TAAAAAAATAAATAAAG 425

RESULT 13
AAZ97059/c
ID AAZ97059 standard; CDNA; 1423 BP.
XX
XX AAZ97059;
AC
XX 19-APR-2000 (first entry)
DT
XX
XX Human secreted protein gene 41 CDNA clone HSIAC80, SEQ ID NO:51.
DE
XX
XX Human; secreted protein; cancer; tumour; developmental abnormality;
KM foetal deficiency; blood disorder; immune system disorder; inflammation;
KM autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
KM schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
KM atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
KM digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
KM therapy; chromosome 12; ds.
XX
XX Homo sapiens.
OS
XX
XX W09966041-A1.
XX
XX 23-DEC-1999.
XX
XX 15-JUN-1999; 99WO-US13418.
XX
XX 16-JUN-1998; 98US-0089507.
XX 16-JUN-1998; 98US-0089508.
XX 16-JUN-1998; 98US-0089509.
XX 16-JUN-1998; 98US-0089510.
XX 22-JUN-1998; 98US-0090112.
XX 22-JUN-1998; 98US-0090113.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Ni J, Rosen CA, Wei Y, Young PE, Florence KA;
PI Soppet DR, Brewer LA, Endress GA, Carter KC, Mucenski M, Edner R;
PI Latleur DM, Olsen HS, Shi Y, Moore PA, Komatsu S G;
XX

```

DR WPI: 2000-106100/09.
 DR P-PSDB; AAY6255.
 PT New isolated human genes and the secreted polypeptides they encode,
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders -
 XX
 PS Claim 1; Page 341; 586pp; English.
 CC AAY627019 to AAY627137 represent 94 isolated human secreted protein genes.
 CC AAY62615 to AAY62633 are the secreted proteins encoded by the 94 human
 CC genes. This sequence represents a fragment of one of the human secreted
 CC proteins. The genes and their corresponding secreted polypeptides are
 CC useful for preventing, treating or ameliorating medical conditions,
 CC e.g., by protein or gene therapy. Also pathological conditions can be
 CC diagnosed by determining the amount of the new polypeptides in a sample
 CC or by determining the presence of mutations in the new genes. Specific
 CC uses are described for each of the 94 genes, based on which tissues they
 CC are most highly expressed in, and include developing products for the
 CC diagnosis or treatment of cancer, tumours, developmental abnormalities
 CC and foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive
 CC disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin
 CC disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney
 CC disorders, digestive/endocrine disorders, infections and AIDS. The
 CC polypeptides are also useful for identifying their binding partners.
 CC The sequences shown in AAY62634 to AAY62655 represent fragments of the
 CC secreted proteins.
 CC
 SO Sequence 1423 BP; 489 A; 215 C; 195 G; 524 T; 0 other;
 QY Query Match 1.0%; Score 21; DB 21; Length 1423;
 Best Local Similarity 100.0%; Pred. No. 9.2;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 635 AAAAATTAAATTTTATTC 655
 ||||||||||||||||||
 523 AAAAATTAAATTTTATTC 503
 RESULT 14
 ID AAF18279 standard; DNA; 1682 BP.
 XX AAF18279;
 AC
 DT 14-MAR-2001 (first entry)
 DE
 XX Lung cancer associated polynucleotide sequence SEQ ID 298.
 KW Human; lung cancer associated protein; neuroprotective; cytostatic;
 KW cardiactive; immunomodulatory; muscular active; vulnarary;
 KW gastrointestinal; nephrotropic; antinefactive; gynecological;
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
 KW proliferative disorder; wound healing; infectious disease; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200055180-A2.
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05918.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.
 XX
 PI Ruben SM;
 DR WPI: 2000-587514/55.

DR P-PSDB; AAB58403.
 XX
 PT Lung cancer associated gene sequences, referred to as lung cancer
 PT antigens, useful for treatment, prevention, and diagnosis of disorders
 PT such as lung cancer -
 XX
 PS Claim 1; Page 757; 1425pp; English.
 CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
 CC associated proteins and polynucleotide sequences, their agonists, and
 CC antagonists may have neuroprotective, cytostatic, cardiactive;
 CC immunomodulatory; muscular active general; vulnarary; gastrointestinal
 CC general; nephrotropic; antinefactive; gynecological; or antibacterial
 CC activity. The invention also includes antibodies specific for the
 CC protein or polynucleotide sequences. The lung cancer associated
 CC polynucleotide sequences may be used for detection of lung cancer,
 CC chromosome identification, as chromosome markers, and for numerous other
 CC diagnostic or research purposes. The proteins may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders. The proteins may also be used in the treatment of wounds and
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
 CC peptide AAB58349 are used in the course of the invention for the
 CC identification and characterisation of the polynucleotide and protein
 CC sequences.
 CC
 SO Sequence 1682 BP; 383 A; 551 C; 430 G; 310 T; 8 other;
 QY Query Match 1.0%; Score 21; DB 21; Length 1682;
 Best Local Similarity 100.0%; Pred. No. 9.2;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1857 GTAAAAAATTAATAAANA 1877
 ||||||||||||||||||
 1537 gtaaaaaataataataaaaa 1557
 RESULT 15
 ID AAZ56739/C standard; CDNA; 1737 BP.
 XX AAZ56739;
 AC
 DT 23-MAR-2000 (first entry)
 DE
 XX Human transmembrane protein HTMPN-42 encoding CDNA.
 KW Human; transmembrane protein; HTMPN; diagnosis; immunospecific;
 KW antiproliferative; neuroprotective; immune disorder;
 KW reproductive disorder; smooth muscle disorder; neurological disorder;
 KW gastrointestinal disorder; developmental disorder;
 KW cell proliferative disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9961471-A2.
 PD 02-DEC-1999.
 XX
 PF 28-MAY-1999; 99WO-US11904.
 XX
 PR 29-MAY-1998; 98US-0087260.
 PR 02-JUL-1998; 98US-0091674.
 PR 02-OCT-1998; 98US-0102954.
 PR 24-NOV-1998; 98US-0109869.
 XX
 PA (INCY-) INCYTE PHARM INC.
 PA
 PI Tang YT, Lal P, Hillman JL, Yue H, Guegler KJ, Corley NC;
 Bandman O, Patterson C, Gorgone GA, Kaser MR, Baughn MR;
 Au-Young J;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2002, 16:46:15 ; Search time 1100.43 seconds
(without alignments)
20409.001 Million cell updates/sec

Title: US-09-636-826-1
Perfect score: 2090
Sequence: 1 GATATCACACATTCGTCCAT.....TCGGCTTCTCGGTGCTGCT 2090

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 11351937 seqs, 5372889281 residues

Word size : 0

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST: *
1: em_estfun: *
2: em_esthm: *
3: em_estln: *
4: em_estom: *
5: em_estpl: *
6: em_estba: *
7: em_estro: *
8: em_estov: *
9: em_htc: *
10: gb_estl: *
11: gb_est2: *
12: gb_htc: *
13: gb_htc: *
14: em_gss_fun: *
15: em_gss_hum: *
16: em_gss_lav: *
17: em_gss_pln: *
18: em_gss_pro: *
19: em_gss_rtd: *
20: em_gss_vtl: *
21: em_gss_other: *

* Pred. No. is the number of results predicted by chance to have a
* score greater than or equal to the score of the result being printed,
* and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	1.5	456	10	AM266586 L0-1455T3
2	32	1.5	496	10	AM266151 L30-2844T
3	32	1.5	548	10	BE131353 L48-1327T
4	32	1.5	564	10	AA962891 L30-421T3
5	32	1.5	568	10	AM266784 L48-122T3
6	32	1.5	569	10	BE037420 MP20H01 M
7	32	1.5	684	11	BF479545 L48-3229T
8	32	1.5	689	11	BF479194 L48-2850T
9	32	1.5	702	10	BE130820 L48-1763T
10	32	1.5	717	11	BF479434 L48-3109T
11	32	1.5	762	10	BE131075 L48-1025T
12	32	1.5	815	10	BE036727 MP04D04 M

13	32	1.5	940	10	BE036821 MP06H04 M
14	31	1.5	911	10	BE034072 MG04G07 M
15	28	1.5	307	11	BE216792 RS136487
16	26	1.2	456	11	D32461 CELK017B8R
17	26	1.2	677	10	A0213676 A0213676
18	25	1.2	894	11	BE675049 602136643
19	24	1.1	590	13	AQ423092 CTRB1-E1-
20	24	1.1	800	13	A2528823 ENTCS3TF
21	24	1.1	852	13	A2683403 ENT1K16TR
22	24	1.1	890	13	A2687045 ENT1L69TF
23	23	1.1	446	10	AL586234 AL586234
24	23	1.1	467	13	A2804361 2M0055H20
25	23	1.1	546	11	BF646364 NF068G02E
26	23	1.1	573	13	A2114192 RPT1-23-4
27	23	1.1	632	11	BF295656 026PbF09
28	23	1.1	635	11	BF298841 068PbF10
29	23	1.1	649	13	A2524829 238PbG12
30	23	1.1	683	13	A2577684 2M0253L10
31	23	1.1	701	13	AG019716 Homo sapi
32	23	1.1	840	13	CNS01EFP
33	22	1.1	208	11	R75193
34	22	1.1	241	10	BS592256
35	22	1.1	308	10	AU062037
36	22	1.1	333	10	AV094649
37	22	1.1	438	10	AM840674 RC4-CN00
38	22	1.1	489	11	BS347359 da95606.
39	22	1.1	511	11	C23717
40	22	1.1	518	10	AU052004
41	22	1.1	543	10	AU155621
42	22	1.1	565	13	A2840399 2M0136M19
43	22	1.1	574	10	A1641746 fc23612 x
44	22	1.1	611	13	BE114399 RPT1-24-2
45	22	1.1	624	13	AZ502227 IM0341I02

ALIGNMENTS

RESULT 1
AM266586 456 bp mRNA EST 20-FEB-2001
LOCUS L0-1455T3 Ice plant Lambda Uni-Zap XR expression library, 0 hours
DEFINITION NaCl treatment Mesembryanthemum crystallinum cDNA clone L0-1455 5',
mRNA sequence.

ACCESSION AM266586 GI:6647072
VERSION
KEYWORDS
SOURCE
ORGANISM
common ice plant.
Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
Caryophyllidae; Caryophyllales; Alzooceae; Mesembryanthemum.

REFERENCE
AUTHORS Cushman, J.C.
TITLE An expressed sequence tag database for the common ice plant,
Mesembryanthemum crystallinum
JOURNAL Unpublished (1997)
COMMENT Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR Primers
FORWARD: T7
BACKWARD: T3
Plate: L0-15 row: E column: 11
Seq primer: T3
High quality sequence stop: 350.
Location/Qualifiers
1. 456
/organism="Mesembryanthemum crystallinum"

FEATURES
source

```

/db_xref="taxon:3544"
/clone_lib="10-1455"
/clone_lib="ice plant lambda uni-zap XR expression library
  0 hours NaCl treatment"
/tissue_type="leaf"
/dev_stage="Six week old"
/Note="Vector: Lambda Uni-zap XR, Bluescript SK-; Site_1:
ECORI; Site_2: XhoI"
BASE COUNT      79 a      102 c      117 g      158 t
ORIGIN

Query Match      1.5%; Score 32; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0091;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2046 AACGGCGATGAACGGCGCGCTTCTTGGGCTT 2077
|||||
Db 79 AACGGCGATGAACGGCGCGCTTCTTGGGCTT 110

RESULT 2
AW266151      496 bp      mRNA      EST      20-FEB-2001
LOCUS      L30-2844T3 Ice plant lambda Uni-zap XR expression library, 30 hours
DEFINITION      NaCl treatment Mesembryanthemum crystallinum cDNA clone L30-2844
5', mRNA sequence.
ACCESSION      AM266151
VERSION      AM266151.1 GI:6644838
KEYWORDS      EST.
SOURCE      Common ice plant.
ORGANISM      Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Caryophyllales; Alzooceae; Mesembryanthemum.
1 (bases 1 to 496)
REFERENCE      Cushman, J.C.
AUTHORS      An expressed sequence tag database for the common ice plant,
TITLE      Mesembryanthemum crystallinum
JOURNAL      Unpublished (1997)
COMMENT      Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR Primers
FORWARD: T7
BACKWARD: T3
Plate: L30-29 row: D column: 12
Seq primer: T3
High quality sequence stop: 350.
FEATURES
source      Location/Qualifiers
1. 496
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/clone_lib="10-2844"
/clone_lib="ice plant lambda Uni-zap XR expression library
  30 hours NaCl treatment"
/tissue_type="leaf, 30 h 0.4M NaCl"
/dev_stage="Six week old"
/Note="Vector: Lambda Uni-zap XR, Bluescript SK-; Site_1:
ECORI; Site_2: XhoI"
BASE COUNT      107 a      110 c      107 g      172 t
ORIGIN

Query Match      1.5%; Score 32; DB 10; Length 496;
Best Local Similarity 100.0%; Pred. No. 0.0088;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2046 AACGGCGATGAACGGCGCGCTTCTTGGGCTT 2077

```

```

|||||
Db 355 AACGGCGATGAACGGCGCGCTTCTTGGGCTT 386

RESULT 3
BE131353      548 bp      mRNA      EST      20-FEB-2001
LOCUS      L48-1327T3 Ice plant lambda Uni-zap XR expression library, 48 hours
DEFINITION      NaCl treatment Mesembryanthemum crystallinum cDNA clone L48-1327,
mRNA sequence.
ACCESSION      BE131353
VERSION      BE131353.1 GI:8578716
KEYWORDS      EST.
SOURCE      Common ice plant.
ORGANISM      Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Caryophyllales; Alzooceae; Mesembryanthemum.
1 (bases 1 to 548)
REFERENCE      Cushman, J.C.
AUTHORS      An expressed sequence tag database for the common ice plant,
TITLE      Mesembryanthemum crystallinum
JOURNAL      Unpublished (1997)
COMMENT      Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR Primers
FORWARD: T7
BACKWARD: T3
Plate: L48-14 row: C column: 3
Seq primer: T3
High quality sequence stop: 350
POLYA+NO.
FEATURES
source      Location/Qualifiers
1. 548
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/clone_lib="L48-1327"
/clone_lib="ice plant lambda Uni-zap XR expression library
  48 hours NaCl treatment"
/tissue_type="leaf, 48 h 0.4M NaCl"
/dev_stage="Six week old"
/Note="Vector: Lambda Uni-zap XR, Bluescript SK-; Site_1:
ECORI; Site_2: XhoI"
BASE COUNT      89 a      126 c      137 g      196 t
ORIGIN

Query Match      1.5%; Score 32; DB 10; Length 548;
Best Local Similarity 100.0%; Pred. No. 0.0085;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2046 AACGGCGATGAACGGCGCGCTTCTTGGGCTT 2077
|||||
Db 85 AACGGCGATGAACGGCGCGCTTCTTGGGCTT 116

RESULT 4
AA962891      564 bp      mRNA      EST      20-FEB-2001
LOCUS      L30-421T3 Ice plant lambda Uni-zap XR expression library, 30 hours
DEFINITION      NaCl treatment Mesembryanthemum crystallinum cDNA clone L30-421 5'
similar to vacuolar H+-ATPase, 16 kDa proteolipid subunit, mRNA
sequence.
ACCESSION      AA962891
VERSION      AA962891.1 GI:3136364
KEYWORDS      EST.
SOURCE      Common ice plant.
ORGANISM      Mesembryanthemum crystallinum

```


REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
An expressed sequence tag database for the common ice plant,
Mesembryanthemum crystallinum
Unpublished (1997)
Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR Primers
FORWARD: T7
BACKWARD: T3
Plate: L30-5 row: C column: 1
Seq primer: T3
High quality sequence stop: 350.
Location/Qualifiers
1. 564
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/clone="130-421"
/clone_1lb="ice plant Lambda Uni-zap XR expression library
30 hours NaCl treatment"
/tissue_type="leaf, 30 h 0.4M NaCl"
/dev_stage="Six week old"
/note="Vector: Lambda Uni-zap XR, Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI"
BASE COUNT 97 a 138 c 132 g 197 t
ORIGIN

Query Match 1.5%; Score 32; DB 10; Length 564;
Best Local Similarity 100.0%; Pred. No. 0.0084;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2046 AACGGCGATGAACGGCGCGCTTCGCGCTT 2077
|||||
DB 128 AACGGCGATGAACGGCGCGCTTCGCGCTT 159

RESULT 5
AM266784 568 bp mRNA EST 20-FEB-2001
LOCUS L48-122T3 Ice plant Lambda Uni-zap XR expression library, 48 hours
DEFINITION NaCl treatment Mesembryanthemum crystallinum cDNA clone L48-122 5',
mRNA sequence.
ACCESSION AM266784
VERSION AM266784.1 GI:6647219
KEYWORDS EST.
SOURCE common ice plant.
ORGANISM Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllidae; Caryophyllales; Alstroceae; Mesembryanthemum.
1 (bases 1 to 568)
Cushman, J.C.
An expressed sequence tag database for the common ice plant,
Mesembryanthemum crystallinum
Unpublished (1997)
Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR Primers
FORWARD: T7

BACKWARD: T3
Plate: L48-2 row: C column: 2
Seq primer: T3
High quality sequence stop: 350.
Location/Qualifiers
1. 568
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/clone="148-122"
/clone_1lb="ice plant Lambda Uni-zap XR expression library
48 hours NaCl treatment"
/tissue_type="leaf, 48 h 0.4M NaCl"
/dev_stage="Six week old"
/note="Vector: Lambda Uni-zap XR, Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI"
BASE COUNT 98 a 133 c 141 g 196 t
ORIGIN

Query Match 1.5%; Score 32; DB 10; Length 568;
Best Local Similarity 100.0%; Pred. No. 0.0084;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2046 AACGGCGATGAACGGCGCGCTTCGCGCTT 2077
|||||
DB 87 AACGGCGATGAACGGCGCGCTTCGCGCTT 118

RESULT 6
BE037420 569 bp mRNA EST 07-JUN-2000
LOCUS MP20H01 MP Mesembryanthemum crystallinum cDNA 5' similar to
DEFINITION vacuolar atp synthase 16 kd proteolipid subunit, mRNA sequence.
ACCESSION BE037420
VERSION BE037420.1 GI:8332436
KEYWORDS EST.
SOURCE common ice plant.
ORGANISM Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllidae; Caryophyllales; Alstroceae; Mesembryanthemum.
1 (bases 1 to 569)
Bohner, H.J., Borchert, C., Brazille, S., Brooks, J., Eaton, M., Ferreira,
H., Kawasaki, S., McCollough, A., Michalowski, C.B., Palacios, C.,
Scara, G., Wheeler, M. and Zepeda, G.R.
Functional Genomics of Plant Stress Tolerance
Unpublished (2000)
Contact: Michalowski, C.B.
University of Arizona
Bio Sciences West room 513, Tucson, AZ 85721, USA
Tel: 520-621-7982
Fax: 520-621-1697
Email: cm@u.arizona.edu
Insert Length: 1 Std Error: 0.00.
Location/Qualifiers
1. 569
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/clone_1lb="MP"
/tissue_type="apical meristem and leaf primordia"
/dev_stage="6 weeks"
/note="3 d 500mM NaCl"
BASE COUNT 110 a 119 c 149 g 191 t
ORIGIN

Query Match 1.5%; Score 32; DB 10; Length 569;
Best Local Similarity 100.0%; Pred. No. 0.0084;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2046 AACGGCGATGAACGGCGCGCTTCGCGCTT 2077
|||||
DB 55 AACGGCGATGAACGGCGCGCTTCGCGCTT 86

```

RESULT 7
LOCUS BF479545 684 bp mRNA EST 20-FEB-2001
DEFINITION L48-3229t3 Ice plant Lambda uni-zap XR expression library, 48 hours
NaCl treatment Mesembryanthemum crystallinum cDNA clone L48-3229
5', mRNA sequence.
ACCESSION BF479545
VERSION BF479545.1 GI:11550372
KEYWORDS EST.
SOURCE Common ice plant.
ORGANISM Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Caryophyllales; Alzooceae; Mesembryanthemum.
1 (bases 1 to 684)
Cushman, J.C.
An expressed sequence tag database for the common ice plant,
Mesembryanthemum crystallinum
Unpublished (1997)
Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR Primers
FORWARD: T7
BACKWARD: T3
Plate: L48-33 row: C column: 5
Seq primer: T3
High quality sequence stop: 350
POLYA-No.

FEATURES
source Location/Qualifiers
1..684
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/clone="L48-3229"
/clone_1lb="Ice plant Lambda uni-zap XR expression library
, 48 hours NaCl treatment"
/tissue_type="Leaf, 48 h 0.4M NaCl"
/dev_stage="Six week old"
/note="Vector: Lambda Uni-zap XR, Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI"
ECORI: Site_2: XhoI"
123 a 162 c 165 g 234 t

BASE COUNT 123 a 162 c 165 g 234 t

ORIGIN

Query Match 1.5%; Score 32; DB 11; Length 684;
Best Local Similarity 100.0%; Pred. No. 0.0078;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2046 AACGGCGATGAACGGCGCTTCTCGGCTT 2077
|||||
DB 89 AACGGCGATGAACGGCGCTTCTCGGCTT 120

RESULT 8
LOCUS BF479194 689 bp mRNA EST 20-FEB-2001
DEFINITION L48-2850t3 Ice plant Lambda uni-zap XR expression library, 48 hours
NaCl treatment Mesembryanthemum crystallinum cDNA clone L48-2850
5', mRNA sequence.
ACCESSION BF479194
VERSION BF479194.1 GI:11550021
KEYWORDS EST.
SOURCE Common ice plant.
ORGANISM Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Caryophyllales; Alzooceae; Mesembryanthemum.
1 (bases 1 to 702)
Cushman, J.C.
An expressed sequence tag database for the common ice plant,
Mesembryanthemum crystallinum
Unpublished (1997)
Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR Primers
FORWARD: T7
BACKWARD: T3
Plate: L48-18 row: F column: 3

REFERENCE 1 (bases 1 to 689)
AUTHORS Cushman, J.C.
TITLE An expressed sequence tag database for the common ice plant,
Mesembryanthemum crystallinum
JOURNAL Unpublished (1997)
COMMENT Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR Primers
FORWARD: T7
BACKWARD: T3
Plate: L48-29 row: E column: 2
Seq primer: T3
High quality sequence stop: 350
POLYA-No.

FEATURES
source Location/Qualifiers
1..689
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/clone="L48-2850"
/clone_1lb="Ice plant Lambda uni-zap XR expression library
, 48 hours NaCl treatment"
/tissue_type="Leaf, 48 h 0.4M NaCl"
/dev_stage="Six week old"
/note="Vector: Lambda Uni-zap XR, Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI"
ECORI: Site_2: XhoI"
123 a 164 c 160 g 242 t

BASE COUNT 123 a 164 c 160 g 242 t

ORIGIN

Query Match 1.5%; Score 32; DB 11; Length 689;
Best Local Similarity 100.0%; Pred. No. 0.0078;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2046 AACGGCGATGAACGGCGCTTCTCGGCTT 2077
|||||
DB 103 AACGGCGATGAACGGCGCTTCTCGGCTT 134

RESULT 9
LOCUS BE130820 702 bp mRNA EST 20-FEB-2001
DEFINITION L48-1763t3 Ice plant Lambda uni-zap XR expression library, 48 hours
NaCl treatment Mesembryanthemum crystallinum cDNA clone L48-1763
5', mRNA sequence.
ACCESSION BE130820
VERSION BE130820.1 GI:8578209
KEYWORDS EST.
SOURCE Common ice plant.
ORGANISM Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Caryophyllales; Alzooceae; Mesembryanthemum.
1 (bases 1 to 702)
Cushman, J.C.
An expressed sequence tag database for the common ice plant,
Mesembryanthemum crystallinum
Unpublished (1997)
Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR Primers
FORWARD: T7
BACKWARD: T3
Plate: L48-18 row: F column: 3

```

Seq primer: T3
High quality sequence stop: 350

FEATURES
Source
Location/Qualifiers
1..702
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/clone_lib="Ice plant Lambda Uni-Zap XR expression library"
/clone_lib="Ice plant Lambda Uni-Zap XR expression library"
/dev_stage="Six week old"
/tissue_type="leaf, 48 h 0.4M NaCl"
/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT
130 a 157 c 170 g 245 t

ORIGIN

Query Match
Best Local Similarity 1.5%; Score 32; DB 10; Length 702;
Best Local Similarity 100.0%; Pred. No. 0.0077;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2046 AACGGGATGAACGGCGCGCTTCCTCGGCTT 2077
|||||
Db 74 AACGGGATGAACGGCGCGCTTCCTCGGCTT 105

RESULT 10
BF479434 717 bp mRNA EST 20-FEB-2001
LOCUS L48-3109T3 Ice plant Lambda Uni-Zap XR expression library, 48 hours
DEFINITION NaCl treatment Mesembryanthemum crystallinum cDNA clone L48-3109
5', mRNA sequence.
BF479434
BF479434.1 GI:11550261
EST.
SOURCE common ice plant.
ORGANISM Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Caryophyllales; Alzooceae; Mesembryanthemum.
1 (bases 1 to 717)
Cushman, J.C.
An expressed sequence tag database for the common ice plant,
Mesembryanthemum crystallinum
Unpublished (1997)
Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR Primers
FORWARD: T7
BACKWARD: T3
Plate: L48-32 row: A column: 9
Seq primer: T3
High quality sequence stop: 350
POLYA-No.

FEATURES
Source
Location/Qualifiers
1..717
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/clone_lib="Ice plant Lambda Uni-Zap XR expression library"
/clone_lib="Ice plant Lambda Uni-Zap XR expression library"
/dev_stage="Six week old"
/tissue_type="leaf, 48 h 0.4M NaCl"
/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT
134 a 174 c 159 g 250 t

ORIGIN

Query Match
Best Local Similarity 1.5%; Score 32; DB 11; Length 717;
Best Local Similarity 100.0%; Pred. No. 0.0077;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2046 AACGGGATGAACGGCGCGCTTCCTCGGCTT 2077
|||||
Db 149 AACGGGATGAACGGCGCGCTTCCTCGGCTT 180

RESULT 11
BE131075 762 bp mRNA EST 20-FEB-2001
LOCUS L48-1025T3 Ice plant Lambda Uni-Zap XR expression library, 48 hours
DEFINITION NaCl treatment Mesembryanthemum crystallinum cDNA clone L48-1025
5', mRNA sequence.
BE131075
BE131075.1 GI:8578438
EST.
SOURCE common ice plant.
ORGANISM Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Caryophyllales; Alzooceae; Mesembryanthemum.
1 (bases 1 to 762)
Cushman, J.C.
An expressed sequence tag database for the common ice plant,
Mesembryanthemum crystallinum
Unpublished (1997)
Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR Primers
FORWARD: T7
BACKWARD: T3
Plate: L48-11 row: C column: 1
Seq primer: T3
High quality sequence stop: 350
POLYA-No.

FEATURES
Source
Location/Qualifiers
1..762
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/clone_lib="Ice plant Lambda Uni-Zap XR expression library"
/clone_lib="Ice plant Lambda Uni-Zap XR expression library"
/dev_stage="Six week old"
/tissue_type="leaf, 48 h 0.4M NaCl"
/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT
151 a 163 c 209 g 239 t

ORIGIN

Query Match
Best Local Similarity 1.5%; Score 32; DB 10; Length 762;
Best Local Similarity 100.0%; Pred. No. 0.0075;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2046 AACGGGATGAACGGCGCGCTTCCTCGGCTT 2077
|||||
Db 47 AACGGGATGAACGGCGCGCTTCCTCGGCTT 78

RESULT 12
BE036727 815 bp mRNA EST 07-JUN-2000
LOCUS MP04D04 MP Mesembryanthemum crystallinum cDNA 5' similar to
DEFINITION vacuolar atp synthase 16 kd proteolipid subunit, mRNA sequence.

```

ACCESSION      BE036727.1  GI:8331736
VERSION        BE036727.1
KEYWORDS       common ice plant.
SOURCE         Mesembryanthemum crystallinum
ORGANISM       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                Caryophyllales; Caryophyllales; Alzooceae; Mesembryanthemum.
REFERENCE      1 (bases 1 to 815)
AUTHORS        Bohner, H.J., Borchert, C., Brazille, S., Brooks, J., Eaton, M., Ferreira,
                H., Kawasaki, S., McCollough, A., Michalowski, C.B., Palacio, C.,
                Scara, G., Wheeler, M. and Zepeda, G.R.
                Functional Genomics of Plant Stress Tolerance
                Unpublished (2000)
TITLE          Contact: Michalowski, C.B.
JOURNAL        University of Arizona
COMMENT        Bio Sciences West room 513, Tucson, AZ 85721, USA
                Tel: 520-621-7982
                Fax: 520-621-1697
                Email: cbm@u.arizona.edu
                An open reading frame exists.
                Location/Qualifiers
                1..815
                /organism="Mesembryanthemum crystallinum"
                /db_xref="taxon:3544"
                /clone_lib="MP"
                /tissue_type="apical meristem and leaf primordia"
                /dev_stage="6 weeks"
                /note="3 d 500mM NaCl"

BASE COUNT     160 a 183 c 197 g 275 t
ORIGIN

Query Match    1.5%; Score 32; DB 10; Length 815;
Best Local Similarity 100.0%; Pred. No. 0.0073;
Matches 32; Mismatches 0; Indels 0; Gaps 0;

QY 2046 AACGGCGATGAACGGCGCGCTTCTTCGGCTT 2077
      |||||m||||||||||||||||||||||||||
      84 AACGGCGATGAACGGCGCGCTTCTTCGGCTT 115

RESULT 13
BE036821      940 bp  mRNA  EST  07-JUN-2000
LOCUS        MP06H04 MP Mesembryanthemum crystallinum cDNA 5' similar to
DEFINITION   vacuolar atp synthase, mRNA sequence.
ACCESSION    BE036821
KEYWORDS     EST.
SOURCE       common ice plant.
ORGANISM     Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Caryophyllales; Alzooceae; Mesembryanthemum.
REFERENCE    1 (bases 1 to 940)
AUTHORS      Bohner, H.J., Borchert, C., Brazille, S., Brooks, J., Eaton, M., Ferreira,
                H., Kawasaki, S., McCollough, A., Michalowski, C.B., Palacio, C.,
                Scara, G., Wheeler, M. and Zepeda, G.R.
                Functional Genomics of Plant Stress Tolerance
                Unpublished (2000)
TITLE        Contact: Michalowski, C.B.
JOURNAL      University of Arizona
COMMENT      Bio Sciences West room 513, Tucson, AZ 85721, USA
                Tel: 520-621-7982
                Fax: 520-621-1697
                Email: cbm@u.arizona.edu
                An open reading frame exists.
                Location/Qualifiers
                1..940
                /organism="Mesembryanthemum crystallinum"
                /db_xref="taxon:3544"
                /clone_lib="MP"

BASE COUNT     151 a 258 c 236 g 259 t 7 others
ORIGIN

Query Match    1.5%; Score 31; DB 10; Length 911;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2047 ACGCGGATGAACGGCGCGCTTCTTCGGCTT 2077
      |||||m||||||||||||||||||||||||||
      46 ACGCGGATGAACGGCGCGCTTCTTCGGCTT 76

RESULT 15
BG216792/c    307 bp  tRNA  EST  21-APR-2001
LOCUS        RST36487 Atcherys RAGE Library Homo sapiens cDNA, mRNA sequence.
DEFINITION   BG216792
ACCESSION    BG216792.1  GI:13742813
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 307)

Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Dahl,T., Thornton,M., Ramachandran,R., Whittington,J.,
Lerner,L., Krasnoc,D., McElligott,K., Clark,S., Mays,R., Smith,E.,
Velooso,N., Hess,J., Cochran,K., Lo,K., Offenbacher,J., Danzig,J.
and Ducar,M.

TITLE
JOURNAL
COMMENT
Activation of Genome-wide Protein Expression Libraries using Random
Nat. Biotechnol. 19 (5), 440 (2001) In press
Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com

FEATURES
source
1. .307
Location/Qualifiers
High quality sequence stop: 307.

1. .307
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="HT1080"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."
BASE COUNT 108 a 39 c 56 g 104 t
ORIGIN

Query Match 1.3%; Score 28; DB 11; Length 307;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 AAAATTATTTAATATATGAAT 564
|||||
Db 287 AAAATTATTTAATATATGAAT 260

Search completed: April 16, 2002, 19:32:40
Job time: 9985 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2002, 16:32:25 ; Search time 52.84 Seconds
(without alignments)
8957.968 Million cell updates/sec

Title: US-09-636-826-1
Perfect score: 2090
Sequence: 1 GATATCACACATTCTGTCAT.....TCGGCTTCTGGGTGCTGCT 2090

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 351203 seqs, 113238999 residues

Word size : 0

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCtUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfilseq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	1.1	5590	4 US-09-050-159-129	Sequence 129, App
2	21	1.0	1381	2 US-08-454-557C-49	Sequence 49, Appl
3	21	1.0	1381	2 US-08-346-426D-49	Sequence 49, Appl
4	21	1.0	1381	2 US-08-450-673C-49	Sequence 49, Appl
5	21	1.0	1381	5 PCT-US95-17111A-49	Sequence 49, Appl
6	21	1.0	1418	5 PCT-US95-17111A-120	Sequence 120, App
7	21	1.0	1442	2 US-08-454-557C-120	Sequence 120, App
8	21	1.0	1442	2 US-08-340-426D-120	Sequence 120, App
9	21	1.0	1442	2 US-08-450-673C-120	Sequence 120, App
10	20	1.0	404	1 US-08-594-031-154	Sequence 154, App
11	20	1.0	634	1 US-08-450-065-1	Sequence 1, Appl
12	20	1.0	634	1 US-08-450-595-1	Sequence 1, Appl
13	20	1.0	2733	1 US-08-676-967-3	Sequence 3, Appl
14	20	1.0	2733	1 US-08-676-974-3	Sequence 3, Appl
15	20	1.0	2733	2 US-09-098-487-3	Sequence 3, Appl
16	20	1.0	4233	2 US-09-120-513-1	Sequence 1, Appl
17	20	1.0	4233	4 US-09-450-105-1	Sequence 1, Appl
18	19	0.9	1229	2 US-08-975-316-85	Sequence 85, Appl
19	19	0.9	1959	1 US-08-342-411A-3	Sequence 3, Appl
20	19	0.9	1959	5 PCT-US94-12883-4	Sequence 4, Appl
21	19	0.9	1982	4 US-08-747-221B-13	Sequence 13, Appl
22	19	0.9	1982	4 US-08-747-221B-15	Sequence 15, Appl
23	19	0.9	1982	4 US-09-005-051-13	Sequence 13, Appl
24	19	0.9	1982	4 US-09-005-051-15	Sequence 15, Appl
25	19	0.9	2144	3 US-08-747-221B-57	Sequence 57, Appl
26	19	0.9	2144	3 US-08-747-221B-59	Sequence 59, Appl
27	19	0.9	2144	4 US-09-005-051-57	Sequence 57, Appl

28	19	0.9	2144	4 US-09-005-051-59	Sequence 59, Appl
29	19	0.9	3528	4 US-08-984-320-2	Sequence 2, Appl
30	19	0.9	3528	4 US-08-487-087A-2	Sequence 2, Appl
31	19	0.9	4383	6 5175095-4	Patent No.5175095
32	19	0.9	12565	4 US-09-345-217-3	Sequence 3, Appl
33	19	0.9	19124	2 US-08-487-826B-13	Sequence 13, Appl
34	18	0.9	69	1 US-08-146-421-2	Sequence 2, Appl
35	18	0.9	301	4 US-09-053-021-3	Sequence 3, Appl
36	18	0.9	345	4 US-09-053-021-8	Sequence 8, Appl
37	18	0.9	1301	4 US-08-804-166-7	Sequence 7, Appl
38	18	0.9	1301	4 US-08-910-951-7	Sequence 7, Appl
39	18	0.9	1442	1 US-08-468-763-18	Sequence 18, Appl
40	18	0.9	1442	2 US-08-393-996A-18	Sequence 18, Appl
41	18	0.9	1560	1 US-08-194-180-2	Sequence 2, Appl
42	18	0.9	1784	3 US-08-991-426-1	Sequence 1, Appl
43	18	0.9	1784	3 US-09-143-470-1	Sequence 1, Appl
44	18	0.9	1934	4 US-08-776-844-1	Sequence 1, Appl
45	18	0.9	2422	4 US-08-952-365-7	Sequence 7, Appl

ALIGNMENTS

```

Result 1
US-09-050-159-129
; Sequence 129, Application US/09050159A
; Patent No. 6197505
; GENERAL INFORMATION:
; APPLICANT: No. 6197505berg, Lelf T
; APPLICANT: Andersson, Maria K
; APPLICANT: Linstrom, Per H
; TITLE OF INVENTION: METHODS FOR ASSESSING CARDIOVASCULAR STATUS AND
; FILE REFERENCE: 1248/ID042
; CURRENT APPLICATION NUMBER: US/09/050.159A
; EARLIER FILING DATE: 1998-03-27
; EARLIER APPLICATION NUMBER: 60/042.930
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 129
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Angiotensin I converting enzyme, 5' region
US-09-050-159-129

Query Match 1.1%; Score 22; DB 4; Length 5590;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1860 AAAAAATAAATAAAGAA 1861
Db 842 aaaaaataaataaagaaa 863

Result 2
US-08-454-557C-49/c
; Sequence 49, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.

```

COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2500
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 1381 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-49

Query Match 1.0%; Score 21; DB 2; Length 1381;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1858 TAAAAAAATATAATAAAG 1878
DB 450 TAAAAAAATATAATAAAG 430

RESULT 3
US-08-340-426D-49/C
Sequence 49, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2500
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 1381 base pairs
TYPE: nucleic acid

STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-49

Query Match 1.0%; Score 21; DB 2; Length 1381;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1858 TAAAAAAATATAATAAAG 1878
DB 450 TAAAAAAATATAATAAAG 430

RESULT 4
US-08-450-673C-49/C
Sequence 49, Application US/08450673C
Patent No. 5948688
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2500
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 1381 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-49

Query Match 1.0%; Score 21; DB 2; Length 1381;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1858 TAAAAAAATATAATAAAG 1878
DB 450 TAAAAAAATATAATAAAG 430

RESULT 5
PCT-US95-17111A-49/C
Sequence 49, Application PC/TUS9517111A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease

NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 1381 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-49

Query Match 1.0%; Score 21; DB 5; Length 1381;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1858 TAAAAAATAAATAAATAAG 1878
|||||
Db 450 TAAAAAATAAATAAATAAG 430

RESULT 6
PCT-US95-17111A-120/C
Sequence 120, Application PC/TUS9517111A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and
TITLE OF INVENTION: Detection of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994

ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 120:
SEQUENCE CHARACTERISTICS:
LENGTH: 1418 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
FEATURE:
NAME/KEY: CDS
LOCATION: 14..1418
PCT-US95-17111A-120

Query Match 1.0%; Score 21; DB 5; Length 1418;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1858 TAAAAAATAAATAAATAAG 1878
|||||
Db 445 TAAAAAATAAATAAATAAG 425

RESULT 7
US-08-454-557C-120/C
Sequence 120, Application US/0845457C
Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 120:
SEQUENCE CHARACTERISTICS:
LENGTH: 1442 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: both
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 15..1139
US-08-454-557C-120

Query Match 1.0%; Score 21; DB 2; Length 1442;

Best Local Similarity 100.0%; Pred. No. 4.4; Mismatches 0; Indels 0; Gaps 0;

QY 1858 TAAAAAAATATAATAAAG 1878
DB 446 TAAAAAAATATAATAAAG 426

RESULT 8

US-08-340-426D-120/C

Sequence 120; Application US/08340426D

Patent No. 5948634

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection

TITLE OF INVENTION: of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESS:

ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/340,426D

FILING DATE: 14-NOV-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 0609,3840002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 120:

SEQUENCE CHARACTERISTICS:

LENGTH: 1442 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: both

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 15..1139

US-08-340-426D-120

Query Match 1.0%; Score 21; DB 2; Length 1442;

Best Local Similarity 100.0%; Pred. No. 4.4; Mismatches 0; Indels 0; Gaps 0;

QY 1858 TAAAAAAATATAATAAAG 1878
DB 446 TAAAAAAATATAATAAAG 426

RESULT 9

US-08-450-673C-120/C

Sequence 120; Application US/08450673C

Patent No. 5948688

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection

TITLE OF INVENTION: of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESS:

ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/450,673C

FILING DATE: 30-MAY-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 0609,3840004

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 120:

SEQUENCE CHARACTERISTICS:

LENGTH: 1442 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: both

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 15..1139

US-08-450-673C-120

Query Match 1.0%; Score 21; DB 2; Length 1442;

Best Local Similarity 100.0%; Pred. No. 4.4; Mismatches 0; Indels 0; Gaps 0;

QY 1858 TAAAAAAATATAATAAAG 1878
DB 446 TAAAAAAATATAATAAAG 426

RESULT 10

US-08-594-031-154

Sequence 154; Application US/08594031

Patent No. 5783182

GENERAL INFORMATION:

APPLICANT: THOMPSON, Timothy C.

TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES

NUMBER OF SEQUENCES: 175

CORRESPONDENCE ADDRESS:

ADDRESSEE: BAKER & BOTT, L.L.P.

STREET: 1299 Pennsylvania Avenue, N.W.

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20004-2400

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/594,031

FILING DATE: 30-JAN-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/006,838

FILED DATE: 16-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Remenick, James
REGISTRATION NUMBER: 36,902
REFERENCE/DOCKET NUMBER: 0A146-0110
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-639-7700
TELEFAX: 202-639-7890
TELEX:
INFORMATION FOR SEQ ID NO: 154:
SEQUENCE CHARACTERISTICS:
LENGTH: 404 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-594-031-154

Query Match 1.0%; Score 20; DB 1; Length 404;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0;
Indels 0; Gaps 0;

OY 1862 AAAAATAAATAAAGAAA 1881
|||||
DB 260 AAAAATAAATAAAGAAA 279

RESULT 11
US-08-450-065-1/c
; Sequence 1, Application US/08450065
; Patent No. 5798105
; GENERAL INFORMATION:
; APPLICANT: Schoenmakers, Johannes G
; APPLICANT: Koelans, Rudolph NH
; APPLICANT: Moelans, Inge IMD
; TITLE OF INVENTION: No. 5798105el protein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; STREET: Smithline Beecham Corporate Patents -US
; STREET: UW2220, Po Box 1539
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450.065
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 424
; INFORMATION FOR SEQ ID NO: 1:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/949645
; FILING DATE: 04-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Jervis, Herbert H.
; REGISTRATION NUMBER: 31,171
; REFERENCE/DOCKET NUMBER: B2992
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-270-5065
; TELEFAX: 215-270-5090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 634 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: PLASMODIUM
STRAIN: FALCIPARUM
IMMEDIATE SOURCE:
CLONE: 16K
US-08-450-065-1

Query Match 1.0%; Score 20; DB 1; Length 634;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0;
Indels 0; Gaps 0;

OY 1858 TAAAAAATAAATAAATAA 1877
|||||
DB 20 TAAAAAATAAATAAATAA 1

RESULT 12
US-08-450-595-1/c
; Sequence 1, Application US/08450595
; Patent No. 5798106
; GENERAL INFORMATION:
; APPLICANT: Schoenmakers, Johannes G
; APPLICANT: Koelans, Rudolph NH
; APPLICANT: Moelans, Inge IMD
; TITLE OF INVENTION: No. 5798106el protein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: Smithline Beecham Corporate Patents -US
; STREET: UW2220, Po Box 1539
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450.595
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/949645
; FILING DATE: 04-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Jervis, Herbert H.
; REGISTRATION NUMBER: 31,171
; REFERENCE/DOCKET NUMBER: B2992
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-270-5065
; TELEFAX: 215-270-5090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 634 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: PLASMODIUM
; STRAIN: FALCIPARUM
; IMMEDIATE SOURCE:
; CLONE: 16K
US-08-450-595-1

Query Match 1.0%; Score 20; DB 1; Length 634;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1858 TAAAAAAATAAATAAAAA 1877
|||||
DB 20 TAAAAAAATAAATAAAAA 1

RESULT 13
US-08-676-967-3
Sequence 3, Application US/08676967
Patent No. 5747317
GENERAL INFORMATION:
APPLICANT: COLLINS, KATHLEEN
TITLE OF INVENTION: Human Telomerase
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Science & Technology Law Group
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/676,967
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC996-055
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415)343-4342
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2733 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-676-967-3

Query Match 1.0%; Score 20; DB 1; Length 2733;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1321 AATGAGGGAAGAACTGTTT 1340
|||||
DB 1087 AATGAGGGAAGAACTGTTT 1106

RESULT 14
US-08-676-974-3
Sequence 3, Application US/08676974
Patent No. 5770422
GENERAL INFORMATION:
APPLICANT: COLLINS, KATHLEEN
TITLE OF INVENTION: Human Telomerase
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Science & Technology Law Group
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA

COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/676,974
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC996-055
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415)343-4342
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2733 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-676-974-3

Query Match 1.0%; Score 20; DB 1; Length 2733;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1321 AATGAGGGAAGAACTGTTT 1340
|||||
DB 1087 AATGAGGGAAGAACTGTTT 1106

RESULT 15
US-09-098-487-3
Sequence 3, Application US/09098487
Patent No. 5917025
GENERAL INFORMATION:
APPLICANT: COLLINS, Kathleen
TITLE OF INVENTION: Human Telomerase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Science & Technology Law Group
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/098,487
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC996-055
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415)343-4342
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2733 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-098-487-3

Query Match 1.0%; Score 20; DB 2; Length 2733;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1321 AATGAGGGAACCTGTTT 1340
|||||
DB 1087 AATGAGGGAACCTGTTT 1106

Search completed: April 16, 2002, 18:12:29
Job time: 6004 sec